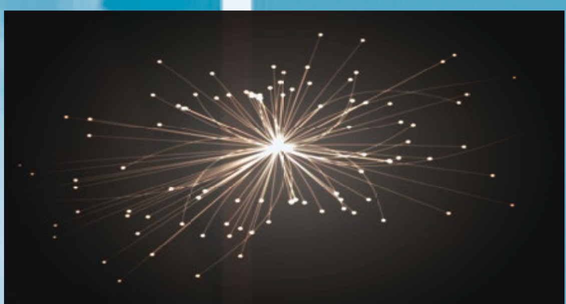
Multiversal Journeys
Founding Editor: Farzad Nekoogar

Yasunori Nomura · Bill Poirier · John Terning

Quantum Physics, Mini Black Holes, and the Multiverse

Debunking Common Misconceptions in Theoretical Physics



Multiversal Journeys

Springer

Multiversal Journeys

More information about this series at http://www.springer.com/series/7919

Yasunori Nomura • Bill Poirier • John Terning

Farzad Nekoogar
Founding Editor

Quantum Physics, Mini Black Holes, and the Multiverse

Debunking Common Misconceptions in Theoretical Physics

Springer

Yasunori Nomura
Department of Physics
University of California
Berkeley, CA, USA
ynomura@berkeley.edu

Bill Poirier
Department of Physics, Department of Chemistry
and Biochemistry
Texas Tech University
Lubbock, TX, USA
Bill.Poirier@ttu.edu

John Terning
Department of Physics
UC Davis
Davis, CA, USA
terning@physics.ucdavis.edu

Founding Editor
Farzad Nekoogar
Multiversal Journeys
17328 Ventura Blvd.
Encino, CA 91316, USA
nekoogar@mvjs.org

Multiversal Journeys™ is a trademark of Farzad Nekoogar and Multiversal Journeys, a 501(c)(3) nonprofit organization

Multiversal Journeys
ISBN 978-3-319-41708-0 ISBN 978-3-319-41709-7 (eBook)
DOI 10.1007/978-3-319-41709-7

Library of Congress Control Number: 2017954954

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Anne Longo – Bill

To my parents – Farzad

To Laura, Jackson, and Maya – John

Preface

A Sun manifested itself as a particle,
And little by little uncovered its face.

All other particles vanished in it;
The world became intoxicated by it and became sober.
Rumi
1207–1273

There is a wide gap between the public understanding of scientific findings and the actual theories developed by scientists in the area of physical sciences. The gap is very noticeable for concepts in theoretical physics and cosmology primarily due to the level of complexity and highly mathematical nature of these theories. Meanwhile, there is a significant amount of misconceptions about theories involved in theoretical physics mainly due to the existence of pseudoscience among some members of the general public.

This book explores, explains, and debunks some common misconceptions about quantum physics, particle physics, spacetime, and multiverse cosmology. It seeks to separate science from pseudoscience. The material is based on a Multiversal Journeys conference at the Lawrence Hall of Science–Berkeley in 2015 with authors of this book as the speakers.

Section one, addressing quantum physics, clarifies what the basic experimental facts imply about the nature of non-locality, the wave function, and what can be measured. It discusses two key quantum experiments: the double-slit experiment and the Einstein-Podolsky-Rosen (EPR) experiment. In both cases, reasoning is by analogy with everyday situations that the reader is already familiar with, while the mathematics is kept to a bare minimum. Interactive web-based animations for the double-slit experiment can be found at www.mvjs.org.

Section two covers misconceptions about the size of elementary particles (no observable size, yet), the structure of atoms (not like mini solar systems), particle colliders (how they are different from but related to microscopes), mini-black holes (why they couldn't have destroyed the Earth when the LHC turned), the Higgs boson destroying the universe (what Stephen Hawking was really talking about), and parallel universes among other topics.

The final section covers multiverse cosmology and related misconceptions, showing that it follows the standard methodology in science: forming a hypothesis about the natural world based on observation (with the help of mathematics) and then looking for evidence that further supports it. The multiverse concept is at a different stage of development when compared to quantum physics and space-time theories. It is a subject of current research activity, rather than a construct of pseudoscience. This section clarifies that the multiverse concept is based on mathematics and that it is a prediction of string theory and eternal inflation. A related documentary about the multiverse is available online at www.mvjs.org.

The material is presented in a layperson-friendly language, followed by additional technical sections which explain basic equations and principles. This feature is very attractive to non-expert readers who nevertheless seek a deeper understanding of the theories and wish to explore beyond just the basic description.

I am grateful to Professor Sean Carroll for reviewing the book. I would like to thank Professor Ken Wharton for his detailed review of the section on quantum physics misconceptions.

I am indebted to the members of our book advisory council: Professor Thomas Buchert, Professor Lawrence M. Krauss, and Professor Mark Trodden, for their suggestions and advice. I would also like to thank the staff of Springer, especially Tom Spicer and Cindy Zitter, for making this project happen.

Finally, I would like to thank the authors of each section, Professor Yasunori Nomura, Professor Bill Poirier, and Professor John Terning for co-authoring the book, and also for their patience throughout the book publication process. Special thanks to Professor Bill Poirier for the macros and formatting style that was used for the book.

Multiversal Journeys Farzad Nekoogar
September 2017

Contents

I Misconceptions in Quantum Physics **1**
Bill Poirier

1 The Singular Quantum **3**

1.1 Why Is Quantum Physics Weird? 3
1.2 What Do the Experts Say? 4
1.3 What Hope Is There for the Rest of Us? 6
1.4 What Material Will Be Covered? 8
1.5 How to Use This Book 10

2 The Bipolar Quantum: Wave-Particle Duality **13**

2.1 Wave, Particle, or "Wavicle"? 13
2.2 Quantum Probability Waves 16
2.3 Heisenberg Uncertainty Principle 20

3 Quantum Soccer: The Double-Slit Experiment **25**

3.1 A Reasonable Doubt 25
3.2 Single-Slit Experiment: Classical 27
3.3 Single-Slit Experiment: Quantum 28
3.4 Double-Slit Experiment: Classical 29
3.5 Double-Slit Experiment: Quantum 30

4 The Plural Quanta: Expanding the Wavefunction **35**

4.1 Many-Particle Wavefunctions . . . 36

4.2 Wavefunction Measurement . . . 38

4.3 Collapse and Schrödinger's Cat . . . 42

5 The "Spooky" Quantum: Nonlocality **47**

5.1 Entanglement and Hidden Variables . . . 47

5.2 Bell's Theorem . . . 48

5.3 EPRB as a Game of Coin Toss: Classical Version . . . 50

5.4 EPRB as a Game of Coin Toss: Quantum Version . . . 54

5.5 Playing the "What If?" Game . . . 56

6 Where the Weird Things Are **61**

6.1 Where Are They? Where Are We? . . . 61

6.2 Interpretations of Quantum Physics . . . 63

6.3 Popular Depictions of Quantum Physics . . . 65

Appendix I: Looking Under the Hood: The Mathematical "Method" Behind the "Madness" **69**

Appendix II: Further Reading **73**

Appendix III: Glossary **75**

II Misconceptions About Particles and Spacetime **79**
John Terning

7 Particles **81**

7.1 What Are Particles? . . . 81

7.2 Atoms . . . 83

7.3 Seeing Atoms . . . 86

8 Mini Black Holes **95**
8.1 Black Holes 95
8.2 Black Holes and Extra Dimensions 98

9 Particle Colliders and the Universe **105**
9.1 The Higgs Boson 105
9.2 Parallel Universes 111

10 Physics in the News **119**
10.1 Cell Phone Radiation 119
10.2 Cold Fusion 123

11 Epilogue **127**
11.1 Spinach: A Cautionary Tale 127
11.2 Moving Forward 128

12 Extra Material: The Equations Behind the Words **131**
12.1 Units and Coordinates 131
12.2 Einstein's Special Theory of Relativity 135
12.3 Quantum Mechanics and Scattering 141
12.4 Gravity 145
12.5 Black Holes 150
12.6 Black Holes in Extra Dimensions 153
12.7 The Large Hadron Collider 155
12.8 Complex Numbers 156
12.9 Group Theory: Rotations and Spinors 157
12.10 The Higgs 161
12.11 Fusion 168
Further Reading 171
More Technical Reading 171
Image Credits 171

III Misconceptions about the Multiverse **175**
Yasunori Nomura

13 The Basic Picture **177**
13.1 Definition 177
13.2 Motivation: Observational 182
13.3 Motivation: Theoretical 185
13.4 New Spacetime Picture 189

14 Demystifying the Multiverse **193**
14.1 Scientific and Conservative 193
14.2 Anthropic “Principle” 196

15 Relation to Observation **199**
Epilogue 203
References and Further Reading 204

Index **207**

Part I: Misconceptions in Quantum Physics

Bill Poirier

Chapter 1

The Singular Quantum

1.1 Why Is Quantum Physics Weird?

If you put the word "quantum" in a talk title at a (non-physics) scientific conference, you will immediately lose the interest of 90% of your audience. However, if you put the same word in a public lecture title, the interest level *increases* tenfold.

What is it about quantum physics that captures the public fascination—yet scares off many professional scientists and engineers—more than just about any other subject? Is quantum physics somehow fundamentally "different" from all other areas of scientific inquiry? Nearly 100 years after its discovery, the "true meaning" of the quantum theory remains as elusive as ever—even as it continues to make the most accurate experimental predictions in all of science.

Quantum predictions of the "fine structure constant" (related to the electron's charge) agree with experiment to better than one part per billion.

This odd dichotomy between ontological ambiguity, on the one hand, and predictive certainty, on the other, does indeed put quantum physics in a rather unique—and perhaps even embarrassing—position among the sciences and also among scien*tists*, who continue to argue with each other about what is really going on. Yet in truth, it seems that we may *never* know what is really going on in the quantum world. This is not simply a matter of not yet having invented the technology needed to perform the right experiments but, rather, appears to be a *fundamental* limitation.

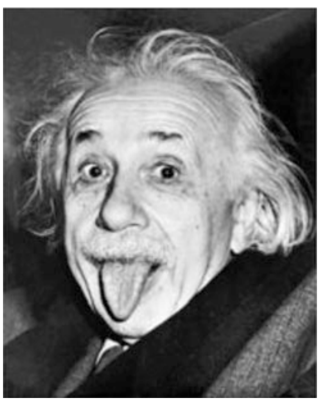
Sometimes they argue more publicly...

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_1

For good or ill, this state of affairs implies that there will likely *always* be debate, dissatisfaction, and disagreement, at least among individuals with a philosophical or foundational bent. On the other hand, in addition to such legitimate controversy, a great many needless misconceptions about quantum physics also abound. This narrative therefore seeks to clarify, if not "debunk," the most common and/or egregious of these misconceptions, without making any dogmatic and/or unprovable claims—e.g., about how the quantum theory is to be interpreted.

1.2 What Do the Experts Say?

Writing on misconceptions about quantum physics is a challenge. In general, misconceptions depend very much on the "beholder," with one person's "obvious fact" being another person's "preposterous fiction." Of course, this is true for any subject, but for quantum physics it is even worse; because the subject is *so* strange and unfamiliar, it is hard *not* to have misconceptions about it.



Even Uncle Albert wasn't always *right...*

We can be forgiven for being slightly confused; despite the tendency of the press to call us "baffled scientists," in truth that's not really news, that's just our natural state! (see Einstein quote below...)

Even the experts are not immune—not even a brilliant one like Albert Einstein. Einstein's own misconceptions about quantum physics are famous, *but*... being Einstein, even his *mis*conceptions pointed the way to important understanding. Alas, most misconceptions that one finds in popular depictions of the subject do not lead to great insight. In any event, we scientists have certainly also contributed our fair share to the general confusion. Indeed, some of our own early misconceptions about quantum physics have continued to propagate even up to the present day.

With those caveats out of the way, what *do* the experts tell us about quantum physics? Here are a few notable quotes, to be discussed in due course:

"If we knew what we were doing, it wouldn't be called 'research'."

—attributed to Albert Einstein

"The 'paradox' is only a conflict between reality and your feeling of what reality ought to be."

—*Richard Feynman*[1]

"I think I can safely say that nobody understands quantum mechanics."

—*Richard Feynman*[2]

"Richard Feynman is probably the most gifted practitioner of quantum mechanics in the first generation to have grown up with it."

—*David Mermin*[3]

Let's start with Einstein. What exactly did he get wrong about quantum physics? Well, he believed that quantum objects possess definite attributes—like position, velocity, shape, etc.—that we are used to associating with *macroscopic* objects, such as baseballs. It was already understood in the early days of the quantum theory that for quantum objects, we cannot actually *observe* all of these attributes simultaneously. However, there was a great debate as to whether the unseen attributes—what we now call "hidden variables"—actually *exist* or not.

Such objects are described by classical *physics—i.e., the familiar, "clockwork" laws of Newton.*

To prove his point, Einstein came up with a famous thought experiment now called the "EPR experiment." The jist of his argument was the following: if quantum theory were correct, and hidden variables do not exist, then the EPR experiment would imply a *crazy* reality—one involving entanglement of faraway particles and "spooky action at a distance." This is what is known as the "EPR paradox." Einstein concluded that quantum theory must therefore be wrong—and in particular, that it must be *incomplete*. However, he also offered a way out: hidden variables would not only "complete" the theory, they would also provide a sensible description of reality.

EPR stands for "Einstein-Podolsky-Rosen," being the three authors on the scientific paper.

[1] From *The Feynman Lectures on Physics*, vol III (Addison-Wesley, 1965).

[2] From *The Character of Physical Law* (British Broadcasting Corporation, 1965).

[3] From *Boojums All the Way Through* (Cambridge University Press, 1990).

The problem with Einstein's reasoning can be summed up in Richard Feynman's first quote above; reality does not always do what we would like it to do. We might be inclined to dismiss the EPR debate as so much metaphysical quibbling, were it not for the work of J. S. Bell. A decade after Einstein's death, Bell raised the stakes by showing that Einstein's "local" hidden variables are incompatible with the *experimental* predictions of quantum theory—which now proved to be not just crazy but **TOTALLY INSANE**.

first by J. F. Clauser and later (and more famously) by A. Aspect...

In fact, EPRB (B for "Bell") experiments have now been performed in the laboratory many times over, starting at the University of California, Berkeley, in the early 1970s.[4] Lo and behold, the utterly insane quantum predictions that Einstein would have found completely inconceivable are in fact *always* observed in the lab, every time. So this makes it official; Einstein's local hidden variables *do not exist*. He got it wrong, but in the end, his ideas nevertheless pointed the way toward important new experiments and insights.

We will discuss the EPRB experiments, action at a distance, and the otherwise total insanity, in Chap. 5.

You don't know Richard Feynman? Really? Google him right now*! Mermin is himself a famous physicist, who wrote* the *classic book on solid state physics.*

The second quote from Richard Feynman above is another fun one—this time, specifically pertaining to quantum mechanics. Just in case you somehow do not know who Richard Feynman is, and question his credentials, the last quote above is by David Mermin, describing Feynman's credentials. In a nutshell, Feynman is the guy who arguably understands quantum mechanics better than anyone else.

1.3 What Hope Is There for the Rest of Us?

So, to summarize the situation thus far, regarding quantum mechanics:

1. Einstein got it wrong.
2. Physicists continue to argue about what it means, after nearly 100 years.
3. The best expert in the world seems to tell us it simply cannot be understood.

[4] J. F. Clauser, in *Quantum [Un]speakables: From Bell to Quantum Information*, R. A. Bertlmann and A. Zeilinger, eds. (Springer, 2002).

What hope is there, then, for the rest of us mere mortals? How can I possibly hope to clarify this thorniest of subjects for a non-expert (but presumably bright, knowledgeable, and intellectually curious) reader such as yourself—let alone for my own self?

World's #1 quantum physicist says its hopeless...

It is a real challenge. Last year, I had lunch with Brian Greene, who is famously good at relating complicated scientific ideas such as string theory to a completely general audience. I have a slightly easier task ahead of me, in that this book is pitched at a somewhat higher "intermediate" level. Nevertheless, I took the opportunity to ask Brian how he was able to do what he does. I paraphrase, but his response was something like the following:

"Explaining string theory is easy; everybody at least understands the concept of a string. Quantum physics, on the other hand, is the hardest."

—Brian Greene

Given the challenge at hand, I am not going to try to make you an expert, and I am certainly not going to attempt to explain away *all* of the mystery of quantum physics. I will, however, at least endeavor to inform you on what is definitively known about the subject vs. that which is just speculation or flat out wrong. As a related but independent goal, I also want you to understand clearly the difference between the quantum *experiments* themselves and the *theories* we construct to try to explain and interpret them.

That would be a fool's errand, in my view.

In short, I want you to become a better *quantum consumer*! For better or for worse, there are a lot of quantum "products" on the market these days—both tangible products, as well as in the marketplace of ideas. For instance, you can procure the services of a new age "quantum healer." Or, if you are more of a do-it-yourself type, you can go to Amazon and purchase a "quantum wand"—harnessing the mysterious and healthful powers of zero-point energy for only $42.

Armed with such understanding, you will hopefully be better prepared when you next encounter the subject elsewhere.

Ironically, there is only one purchase and a 0% satisfaction rate—about as "quantum" as you can get!

Perhaps most telling of all, if you do a search on amazon for quantum-themed books, you will find a whopping 37,500 titles for sale. Most of these, I can assure you, are not

On the Web: `www.amazon.com`.

meant for practicing physicists, but for the general public, who seem to have a huge appetite these days for all things quantum. Like it or not, there is a huge quantum market out there that you are being exposed to.

We will speculate on the reasons for this in Sect. 6.3.

You *are* a quantum consumer, and in the same way that you do not need to be an expert in auto mechanics to learn how to buy a good car, you can learn to be a *smart* quantum consumer—someone who understands what they are buying, what to look for in the quantum product, what questions to ask the sales rep, etc. My aim here is to provide you with the tools you need to achieve this—like a kind of *Quantum Consumer Reports*.

1.4 What Material Will Be Covered?

First and foremost, I aim to ground the discussion by focusing on the basic experimental facts, above all else. I will then address what the basic facts concerning quantum physics imply about the nature of:

1. The quantum wavefunction.
2. What can be measured.
3. Quantum nonlocality.

Only then will I discuss the more esoteric aspects—i.e., the headline-grabbing ramifications of the various metaphysical *interpretations* of quantum physics.

In the next few chapters, I will provide an overview of the basic quantum landscape, addressing key issues and misconceptions as I go. I will start with the simplest ideas first and then gradually build up in complexity. Specifically, I will start with wave-particle duality and work up to the Heisenberg uncertainty principle, the double-slit experiment, and Schrödinger's cat. For the most part, the presentation in these sections will not be very technical. However, the level will "ramp up" considerably for the discussions on wavefunction measurement (Sect. 4.2) and the EPRB experiments (Chap. 5). The latter exemplifies that most mysterious of all quantum phenomena, *nonlocality*.

not that there are *any truly simple ideas in quantum physics...*

though I will make every effort to be as accurate as possible...

In keeping with my goal to make you a smart quantum consumer, I will adopt a mostly "no-nonsense" approach, focusing on the less controversial aspects of the subject. Likewise, I will keep the philosophical and historical discussion to a minimum—though it would not be proper, nor even possible, to excise these interesting and important aspects entirely. That said, in the final chapter (Chap. 6), I will engage in a slightly more speculative and philosophical discussion of interpretations (as well as popular depictions), in an attempt to "peek behind the curtain" a bit, to get a sense of what might really be going on back there.

That is to say, those aspects that have been proven in the laboratory beyond any shadow of a doubt.

Historical aspects will be mostly avoided for the simple reason that I am no historian of science. Perhaps a better reason, however, is that the development of quantum physics was quite confused and divisive at its start. Moreover, some of the early missteps are still leading some of us astray to this day. So, rather than hash out what was debated and believed at the time—interesting though it may be—I will focus much more on what is definitively known today, with the benefit of a century's worth of experimental hindsight.

the aforementioned "baffled state"...

The two key quantum experiments that I will describe in detail are:

1. The double-slit experiment.
2. The EPRB experiment.

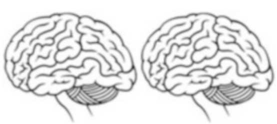
The corresponding chapters—i.e., 3 and 5, respectively—really form the heart of this work. In both cases, I will reason by analogy with familiar situations and everyday objects. For 1, I will use soccer balls kicked down a field. For 2, the analogy is rather different, involving a pair of coin detectors. In somewhat less detail, I will also address more recent experimental and theoretical developments, including: quantum tomography and weak measurement; quantum interference with increasingly large objects; many-worlds interpretation.

1.5 How to Use This Book

i.e., it spans a wide range of abilities. . .

As discussed, this book is intended for an intermediate-level audience, in terms of scientific and mathematical background. The problem is that the word "intermediate" covers a multitude of sins. Then there is the question most dreaded by science author and reader alike: "how much math do I need to know for this?" The answer is, there *will* be some math on this test. However, in order to encompass as broad a swath as possible, much of the conceptual discussion will involve no mathematics at all.

▷▷▷ **Math Alert!** When some math *is* required, I will usually try to warn you with these "brain" boxes. Even then, I have tried to keep the math to a bare minimum and the level no harder than (and, in one case, very similar to) that of a Sudoku puzzle.



▷▷▷ **Math Alert! *double trouble!!*** In rare cases, when I have to resort to more advanced mathematics, I will signify this using these *double*-brain areas. The entire Appendix I, for example, is one great double brain, wherein the actual working mathematical equations that underlie all of quantum physics are revealed!

▷▷▷ **Misconception !!** As per the title of this book, I also rely on the device of pointing out popular misconceptions explicitly and highlighting these within their own special boxes, such as this one. In addition to "Misconception" boxes. . .

▷▷▷ **More Accurate !!** . . . there are also "More Accurate" description boxes, such as this one, wherein I attempt to replace a misconception with a "better" (albeit certainly not perfect) description.

▷▷▷ **Lesson:** Finally, when there are lessons to be learned, I will summarize them in "Lesson" boxes such as this one. Particularly in the arena of quantum physics, the hope is that this approach will serve as an effective means of communication and instruction.

No one "just" reads books any more—or "just" goes to class—or "just" watches TV. We recognize that ours is an age where everyone is glued to their devices at all times. Accordingly, I will occasionally exhort you to do a Google search or direct you to a particular website. The latter I will do using marginal notes with an image of a globe and the words "***On the Web***."

as on p. 6...

As an example, now is a good time to advertise the companion website for this book, whose url is provided in the margin just next to this text. Here, you can find interactive web-based animations, e.g., for the double-slit and EPRB experiments, enabling you to simulate your own virtual quantum experiments. Even though you haven't read those chapters yet, go ahead and have at it. Statistically, we know it's going to happen anyway, and besides, the book will still be here when you get back.

On the Web: `www.mvjs.org`.

Chapter 2

The Bipolar Quantum: Wave-Particle Duality

2.1 Wave, Particle, or “Wavicle”?

It’s high time that we got on to debunking some misconceptions. Let’s start with the wave-particle duality. If you know anything at all about quantum physics, you probably know that tiny objects have both wave-like and particle-like attributes. Sometimes this fact is stated a little too strongly as follows:

▷▷▷ **Misconception !!** Tiny objects are both waves and particles at the same time.

This is a myth because particles and waves are completely different and mutually exclusive things—*at least in the conventional, classical sense in which these concepts are usually understood.* In particular, particles are always “localized” (located at a single point) in space, whereas waves are always “delocalized” and continuous (spread out over many points in space). So how can any single object be both of these things at the same time?

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*, Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_2

It is more accurate to say the following:

▷▷▷ **More Accurate !!** Tiny objects *sometimes* behave like waves and *other times* like particles.

While still weird, this statement at least makes logical sense. Quantum objects exhibit wavelike and particle-like aspects, *under different circumstances*, but they cannot actually *be* both things at once. That would be a logical fallacy. In truth, tiny objects are *neither* waves *nor* particles, nor even really a hybrid of the two, but something quite different that neither the wave nor the particle concept can adequately describe. In 1928, Sir Arthur Eddington coined the term "wavicle" to remind us of this inherent schizophrenia of quantum objects—which are not quite particle yet not quite wave. Still, like many preconceived notions, these concepts are so fundamental and familiar to us that we are loathe to give them up, even today.

Eddington earned a knighthood after confirming general relativity through his famous 1919 eclipse observation. He also declared that the quantum world consists of "mind stuff."

What is the smart quantum consumer to take away from all of this? I am smelling our first lesson coming on, and it is an extremely important one:

▷▷▷ **Lesson:** Accept the weird if you have to, but don't accept the logically impossible.

Physics may be alien, uncanny, bizarre, and counterintuitive, but it cannot be inconsistent—at least not insofar as experimental observation is concerned.

By the way, speaking of consistency, I will not always be very precise with language. For example, I will routinely refer to the tiny objects in question as "quantum particles," even when they are acting like waves rather than particles. Also, I will be rather vague about what "tiny" means, exactly. As a rule of thumb, this means objects on the nanoscale or smaller, such as molecules, atoms, or subatomic particles. In truth, however, the effective scale at which things start to become "quantum" varies signifi-

This is standard practice, so don't get confused.

i.e., 1 nanometer, 10^{-9} m.

cantly with context. For the most part, we will not worry too much about where and how this transition occurs, exactly.

except when discussing the Schrödinger cat dilemma in Sect. 4.3...

The More Accurate description above may sound rather innocuous, and it can sometimes lead people to go too far in *that* direction, via the following:

▷▷▷ **Misconception!!** There is nothing mysterious or weird about quantum wave-particle duality; these are just two mutually compatible sides of the same coin.

A classic example is that famous popularizer of science, Isaac Asimov, who wrote[1]:

“A man may have many aspects: husband, father, friend, businessman... You would not expect him to exhibit his husbandly behavior toward a customer or his business-like behavior towards his wife.”

—*Isaac Asimov*

With all due respect to Uncle Isaac, this is not really correct either. *Something* very peculiar really *is* going on in quantum physics. There is no question about this; it is a basic experimental fact, as we shall see. In any event, there is more to wave-particle duality than the simple notion that one and the same object exhibits different characteristics under different circumstances, although that is certainly a part of it.

at least insofar as our classically intuitive concepts are concerned...

▷▷▷ **More Accurate!!** Wave-particle duality is *very* weird.

▷▷▷ **Lesson:** Accept the weird in quantum physics. Don’t accept the notion that the weirdness can be entirely “explained away.” It cannot be—at least not within the context of our current understanding.

[1] from *The New Intelligent Man’s Guide to Science* (Basic Books, New York, 1965).

So what have we learned here? On the one hand, people sometimes try to attribute *more* weirdness to quantum physics than is really there, by imagining that it can support logical contradictions. On the other hand, others try to explain away *all* of the inherent weirdness and mystery. In my view, both tendencies are problematic and should be avoided.

2.2 Quantum Probability Waves

What makes quantum wave-particle duality so strange is the manner in which this duality manifests—and also, the manner in which the quantum wave is interpreted. As a bit of technical jargon, the wave is described by something called the *wavefunction*—a theoretical construct whose purpose is to tell us what the particle is doing. The wavefunction can never be measured directly, which may already seem a bit strange. The *really* strange thing about it, though, is that it behaves completely differently, depending on whether or not the particle it describes is being *observed*.

although see Sect. 4.2. . .

How does it know?

When the particle is *not* being observed (i.e., measured), it behaves like a wave. The wavefunction then describes a delocalized "wave of probability," spanning all possible particle positions throughout space. The particle acts as if it is somehow able to explore all possibilities simultaneously—in this case, all possible positions in space. We thus say that the particle is in a "superposition" of position states. It has become wave*like*, although it is not necessarily true that the particle actually *becomes* a wave. Over time, the wavefunction evolves according to well-prescribed rules that tell us very precisely what happens to the set of all possible particle states. These rules, moreover, are completely deterministic.

We do not know for sure, precisely because the particle is not *being monitored!*

The technical name for these rules is the "time-dependent Schrödinger equation" (see Appendix I).

When the particle *is* observed, the wavefunction *also* tells us what happens—but in a *very* different way. First, it is an experimental fact that a quantum particle is *always* observed as a *particle*—never as a wave nor as multiple "copies." So, when a measurement of a particle is made, and a definite position revealed, the wavefunction is said to "collapse" to the observed position. This is a

The wave of infinite possibilities "collapses" down to just one.

random process, in that the particular observed position cannot be predicted in advance. However, it is not *completely* random either, meaning that the particle is more likely to show up in some regions of space than in others. How do we know where it is most likely to turn up? The wavefunction itself provides us with these probabilities.

"Random" implies that the wavefunction collapse is not *deterministic.*

with exceedingly *high accuracy...*

The wavefunction, therefore, plays an interesting dual role. On the one hand, it tells us how the particle explores all of its possibilities when we are *not* looking at it; on the other hand, it tells us where it will likely turn up when we *do* look. Of the two, it is the second "collapse" role that causes most of the trouble.

This is reminiscent of the two very different roles of DNA—i.e., transcription and translation.

Since only particles can be measured—and not their underlying wavefunctions—how do we know for sure that wavefunctions really exist? The truth is, we don't. It may well be that wavefunctions are "real"; then again, there are alternative theoretical constructs that also do the same job just as nicely. In any case, the wavefunction construct is certainly a useful one, providing "standard" quantum theory with the means to make practical, concrete, quantitative predictions.

see Sect. 4.2.

The above teaches us another important lesson:

▷▷▷ **Lesson:** Quantum physics is about exploring all possibilities. When they are not being observed, quantum particles explore all possible states available to them, such as positions in space.

This is true, but it needs to be interpreted correctly. Though "all possibilities" are explored, this occurs only within the confines of certain well-prescribed and highly constrained rules. It definitely does *not* mean that "anything goes," which is simply not true.

▷▷▷ **Misconception !!** Quantum physics is *so* weird that "anything can happen" and nothing is certain.

▷▷▷ **More Accurate !!** Quantum physics is *the* most reliably accurate predictive scientific theory ever devised.

It is worth pausing to consider what the More Accurate description above really means. It means that there is, in fact, very *little* wiggle room for a quantum theory. This is one of the most important and basic things that many people do not understand. They think that scientists use quantum mechanics as an excuse to indulge in all sorts of wild speculation. Sometimes scientists *do* speculate—especially in new areas where there is not much data. That's fine; you have to get started with some kind of hypothesis.

But quantum physics is *well beyond* that point! Any crazy, wild-eyed new quantum theory that someone comes up with now had better agree with the reams of experimental data that have been amassed over the last 100 years. Thus do we come upon another important lesson for the smart quantum consumer:

▷▷▷ **Lesson:** Any physical theory *must agree* with all relevant experimental observations, or it has to be discarded.

There is little disagreement anymore, about what is observed in the quantum laboratory; for the most part, this is not where the debate is. On the other hand, where there still *is* a bit of wiggle room—and likely always will be—is in the *interpretation* of the theory.

All of that said, it is weird enough that quantum particles get to explore all possible positions when they are not being watched. This causes enough headaches as it is—both conceptual and also computational, if one is trying to simulate how these things behave on a computer. In particular, one has to worry about things like this happening[2]:

which happens to be what I do for a living; the last *thing I want is to have to open the door to even* more *bizarre possibilities. . .*

[2] with apologies to Charles Addams and *The New York Times* (December 3, 2006).

This sort of thing is not a "fluke" or rare occurrence in quantum physics; it happens all the time.

We will see something very similar in the double-slit experiment (Chap. 3).

I conclude this section with a homage to the late great Yogi Berra. Yogi Berra was an American icon. Yogi Berra was a great baseball player and a great wit. Yogi Berra also had many wonderful quotes that just happen to describe quantum mechanics *perfectly*. In particular, there is a great one involving the word "fork." Unfortunately, it is difficult to get these quotes to appear in print for a reasonable price. A shame...

On the Web: `www.google.com`.

Yogi Berra

Yogi Bear

Separated at Birth?

2.3 Heisenberg Uncertainty Principle

The notion of a *Principle* was big in the early days of quantum physics. Everybody who was somebody had to have one. Thus, you have your "Pauli Exclusion Principle," your "Principle of Superposition," your "Principle of Complementarity," and the "Born Rule," which is actually a Principle. But for sure, without a doubt, the hands-down Grand-Daddy of all the Grand Quantum Principles has got to be the "Heisenberg Uncertainty Principle" (HUP).

The Born Rule is what tells us how to compute quantum measurement probabilities from the wavefunction.

Accordingly, it is perhaps not surprising that no aspect of quantum physics has been more *mis*understood than the HUP. So let's start straightaway with what the HUP is *not*. It is most definitely *not* this oft-heard statement:

▷▷▷ **Misconception !!** The Heisenberg Uncertainty Principle states that "no measurement can be made of a quantum particle without affecting that particle."

While the statement above (in quotes) is not wrong, it is in fact *always* true in physics and has nothing to do with quantum mechanics per se.

This latter statement is sometimes referred to as the "observer effect."

It is certainly true that the HUP implies that quantum measurement is very *different* from classical measurement. For instance, the latter can always be devised so as to have *minimal* impact on subsequent measurements, at least in principle. For many quantum measurements, this is not possible even *in principle*; the system is necessarily greatly altered as a result of measurement, regardless of the particular measurement outcome. I will address quantum measurement issues more fully in Chap. 4.

It is better to say that the system as a whole*—including both observer and observed—experiences the change, rather than just the measured particle itself.*

For now, the main point is that the HUP is a *limiting* principle; it limits the extent of knowledge that we are allowed to have about quantum systems—at least, in comparison with classical systems. Moreover, it is a *fundamental* property of all quantum systems and measurement devices—rather than some lament about the limits of present-day experimental apparatuses. No matter how technology progresses, we will *never* be able to get around the HUP.[3]

although in a somewhat technical sense, this can be, and already has been, accomplished (see footnote)...

So what actually *is* the HUP then? It is this:

▷▷▷ **More Accurate !!** There is a fundamental limit to the precision with which both a particle's *position* and its *velocity* can be determined simultaneously.

The more carefully we measure a particle's position, the less well we can know its velocity (and vice versa). It can be instructive to think of this state of affairs as being where quantum wave behavior comes from. Consider that any particle—in order to remain a "particle" (in the classical sense) over time—must possess definite values for *both* attributes, position, *and* velocity. Conversely, in whatever sense the HUP denies this possibility, it gives rise to a delocalized, *wavelike* aspect of quantum particles.

assuming that the particle even has *a well-defined velocity, which may not be the case...*

i.e., localized in space...

We can see this explicitly in the following example. Suppose you watch the arc of a baseball, after it has been thrown. At each point in time along its trajectory, the baseball has both a definite position and velocity, as well as other attributes that you can observe. But suppose that at a given point in time, only the *position* of the baseball is well determined and *not* its velocity. Then, in the next instant of time, having a range of velocities available to it, the baseball would effectively come to occupy a *range* of different positions simultaneously. In other words, it would become a wave of superposed position states.

thrown by a "classic," say, Sandy Koufax...

now this *is clearly a baseball thrown by Yogi Berra...*

Now, we can always measure a quantum particle's position, and when we do so, we always find it to be in a certain place—i.e., we find that it behaves like a particle. But it does so only *instantaneously*. As soon as we look away, the HUP dictates that the particle must start behaving like a wave again—precisely *because* its velocity is not well determined.

[3] http://www.emqm15.org/presentations/speaker-presentations/yuji-hasegawa/

▷▷▷ **Lesson:** We cannot "see" all classical attributes of quantum particles at once, the way we can for classical particles. At best, the unseen attributes have undetermined values; at worst, they do not exist at all.

The above lesson exemplifies a *very* key point, probably the most important of all—which is why you already saw it in Sect. 1.2 and will see it again later on. Let me remind you, also, that there really is *no* debate on this question, which has been confirmed by experiment many times over. When we observe quantum systems, we are—in comparison with classical observation—forever "partially blind." This has nothing fundamentally to do with the "smallness" of quantum particles; as my coauthor John Terning is fond of saying, a giant microscope will not help us to "see" the quantum world any better.

No experiment has ever measured precisely a particle's position and velocity at the same time.

Consult John Terning's part of this book for an excellent description of how quanta manifest *in practice, in electrons, atoms, photons, and so forth.*

▷▷▷ **Lesson:** Not everything in science can be completely understood! In particular, we can *predict* very accurately what happens in quantum experiments, but we cannot necessarily *explain how* it happens.

On the other hand, a much more subtle question is whether or not the unobserved attributes—known as "hidden variables"—actually *exist* for quantum particles, as Einstein believed. This question is still not completely answered. However, one thing is certain: if they do exist, then they must behave very strangely.

In jargon, they must be global *hidden variables. More on that in Sect. 5.2.*

All of this serves to pose some interesting philosophical questions. In particular, the question of whether something exists if we cannot see it presents a profound epistemological quandary. In effect, quantum physics and the HUP have put the age-old "If a tree falls in the forest?" conundrum on steroids. . .

"I like to think that the moon is there even if I am not looking at it"—another great Einstein quote. . .

Werner Heisenberg, Uncertainty Principle *wunderkind*, Director Max Planck Institute for Physics Göttingen, Nobel Laureate.

"Heisenberg," blue meth *wunderchemiker*, California Institute of Technology, Cofounder Gray Matter Technologies.

Separated at Birth?

Chapter 3

Quantum Soccer: The Double-Slit Experiment

3.1 A Reasonable Doubt

Based on the description in Sect. 2.3, a perceptive reader might well question the reality of quantum probability waves by posing something like the following scenario. Suppose one observes a *classical* object, such as a bowling ball flying down a lane. Halfway to the pins, the observer shuts his or her eyes and keeps them shut until the ball reaches the end of the lane. Can the observer predict with certainty which path will be taken and which pins will be knocked over? Certainly not. Due to his or her imperfect knowledge, it is *as if* the bowling ball has become a "wave of probability" across a superposition of possible outcomes, only one of which will eventually be realized.

In Sects. 1.2 and 2.3, it was a baseball*, and it will soon become a* soccer *ball.*

Of course, there is nothing at all mysterious going on here. In reality, the ball follows a *single* path only; the "wave" simply reflects the observer's ignorance about which path that is and does not influence the actual state of the ball itself. We call this kind of wave a wave of "classical statistical probability."

Might something similar be going on in quantum physics? There are at least two reasons why the answer is a resounding "no." The first we have already addressed: the

People like A. Einstein certainly hoped so.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*, Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_3

HUP is a *fundamental* limitation that cannot be ameliorated through improved observation. Classical probability waves, on the other hand, allow for *any* degree of observer certainty, at least in principle. Indeed, for the above bowling ball example, all the observer needs to do to remove all uncertainty is to keep his or her eyes open the whole time.

an act which still has no effect on the ball...

The second reason is a bit more subtle but also much more weird. In classical statistical theory, the different parts of the probability wave behave independently. After all, only one path is the "real" path actually followed by the classical object, which cannot be influenced by the "paths not taken." A quantum probability wave, on the other hand, behaves much more like a *tangible* wave, in the sense that different parts of the wave *do* interact with each other.

For instance, when two parts of a tangible wave collide, the result is often *wave interference*—a pattern of alternating "high" and "low" fringes, with a spacing that relates directly to the wavelength. Classical probability waves *never* manifest interference—behaving instead more like stereotypical "bell curves."

or "bright" and "dark," in the case of light waves...

though "never" is a slightly dangerous word...

Guess what quantum probability waves do?

The double-slit experiment is extremely important, because it provides a direct and dramatic example of quantum wave interference, *in the laboratory.* It contains, as R. Feynman once stated, "the *only* quantum mystery." It is conceptually very simple yet leads immediately to a seeming paradox that defies all conventional explanation. Most importantly for our purposes, however, it provides *direct experimental proof* that *something* very weird really is going on.

Although the double-slit experiment is a staple in virtually all elementary discussions of quantum physics, I will address it in a slightly different way than most. In particular, I will break it down into a sequence of *four* separate experiments—in an attempt to pin down exactly where the weirdness is coming from. As we will see, however, the weirdness cannot be pinned down...

3.2 Single-Slit Experiment: Classical

The first experiment is so simple that you can easily reproduce it yourself. This is the *classical* single-slit experiment. A soccer player stands at one end of a field, firing soccer balls at a wall with a single vertical hole or slit in it. For each soccer ball that is kicked, we monitor and record what happens to it over time. Then, we repeat the experiment, many many times over. Note that *only one ball is ever in play at a time!* This is very important to keep in mind.

In the lecture that preceded this book, it was Carli Lloyd, but it can be whomever you wish.

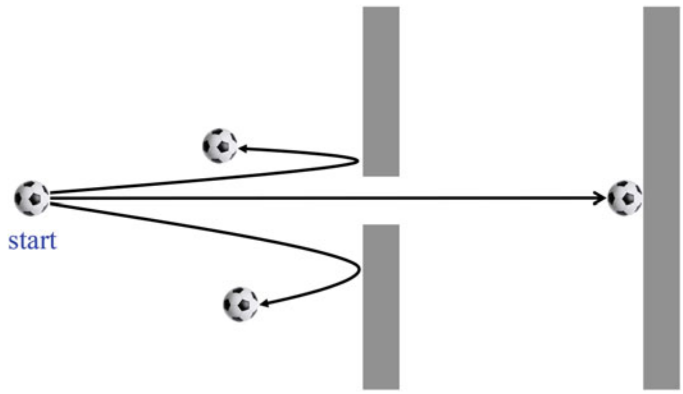


Fig. 3.1 **Classical single-slit experiment.** Soccer balls are fired one at a time toward a wall with a slit; many do not make it through.

So what happens? Many of the balls will simply bounce off the first wall; but some will go through the slit and continue on to hit a *second* wall, placed behind the first, as indicated in Fig. 3.1 above. The "goal" of the experiment is to monitor—for all of the balls that *do* make it through the slit—where they land on the second wall. We can expect that most of the balls will land directly behind the slit, with a few landing to either side. Very occasionally, there might also be an outlier, a bit further out. After many many such repeated observations, we build up a classical probability distribution that looks like a bell curve:

no pun intended...

Fig. 3.2 **Classical single-slit experiment.** Those balls that *do* make it through the slit form a bell-curve distribution on the far wall.

3.3 Single-Slit Experiment: Quantum

This one may be a bit harder for you to try at home...

Next, we consider the *quantum* version of the single-slit experiment. Imagine what happens when you shrink the soccer balls down to the nanoscale. The balls now become quantum particles and as such can no longer be "seen" during the course of their journey. As in the classical case, each nano-soccer ball *starts out* as a particle, occupying a definite position on the left end of the field, as indicated in Fig. 3.3. Once kicked, however, we cannot say precisely what happens. Something wavelike moves along, evidently goes through the slit, and encounters the far wall. It is then measured, causing it to "collapse" back into a nano-soccer ball—a particle once more, with a definite location.

If you repeat this experiment many times, you will find the same bell-curve probability distribution as in the classical case, i.e., Fig. 3.2. This does not seem so bad. True, we don't know exactly what happens at intermediate times, but the final result is the same as before. We might therefore be tempted to imagine that quantum soccer balls are just like classical ones, after all, with the "hidden" variables behaving just like their classical counterparts.

except for the fact that we cannot see them...

If that were the whole story, then you would have most likely never heard of a small, esoteric branch of science known as "quantum physics," and you would consequently

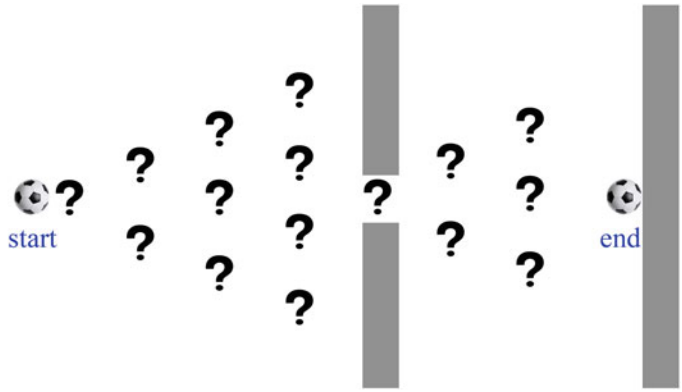


Fig. 3.3 **Quantum single-slit experiment.** Nano-soccer balls are fired towards the slit, and some reappear on the far wall. What happens in between?

not be reading this book. Put another way, there is a reason why one tends not to hear about the "quantum *single*-slit experiment."

except from certain perverse authors...

3.4 Double-Slit Experiment: Classical

Let's scale the soccer balls back up to their usual macroscopic size, and redo the classical experiment—this time with *two* slits, instead of one. Once again, we can follow each soccer ball as it moves along each step of its path to the far wall. We can expect about half of those paths to go through the upper slit and about half through the lower slit.

Note that it makes no difference whether we keep our eyes open or closed...

Since there is only one ball in the air at a time, the paths are completely independent. Each slit on its own therefore behaves as in the single-slit experiment. The resultant double-slit probability distribution is therefore just a combination of two bell curves, one centered behind each slit.

...the final probability distribution is that of Fig. 3.4 in either case.

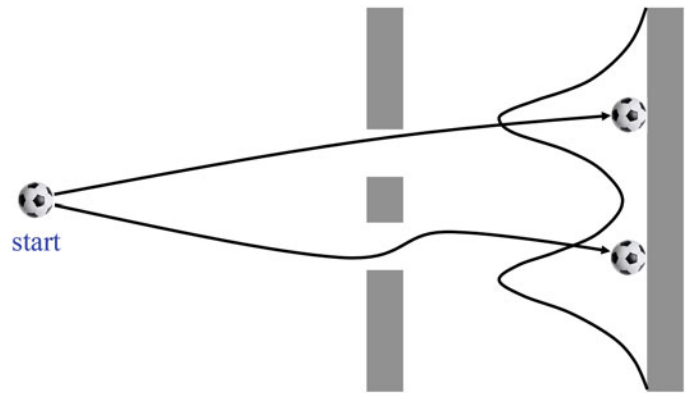


Fig. 3.4 **Classical double-slit experiment.** The probability distribution is just the sum of two bell curves, one for each slit.

3.5 Double-Slit Experiment: Quantum

Now we are ready for the main event, the *quantum* double-slit experiment. Suffice to say, this case behaves nothing like any of the previous three examples. The setup should be straightforward by this stage, so let's jump straight to the punchline. When there are two slits, the quantum probability distribution on the far wall becomes this:

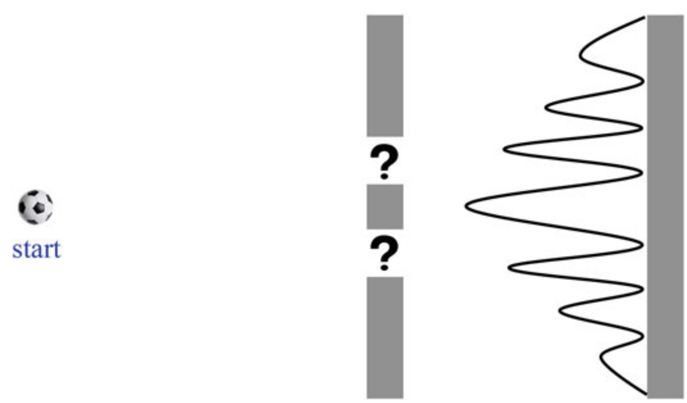


Fig. 3.5 **Quantum double-slit experiment.** What in blazes is going on? *Quantum wave interference*, that's what.

Figure 3.5 represents the fundamental quantum weirdness—something predicted by theory, and confirmed by experiment many many times, yet still very strange to behold. To quote R. Feynman more fully than in Sect. 3.1[1]:

> *"This has in it the heart of quantum mechanics. In reality, it contains the only mystery."*
>
> —*Richard Feynman*

Among other oddities, Fig. 3.5 implies that there are points on the far wall where the nano-soccer ball simply *refuses* to ever land, no matter how many times you kick it. In any event, Fig. 3.5 is an experimental fact, and it is for this reason that we will probably never be able to explain away quantum weirdness completely—no matter how hard we may try.

So how *can* we hope to understand what is going on in Fig. 3.5? Ironically, what makes this pattern so strange is how *familiar* it is—but in the very different context of *tangible* waves, i.e., fluids and light waves and such. This is a *wave interference pattern*, caused by sending the wave through both slits simultaneously and then allowing the two parts to interact.

Figure 3.5 suggests that the nano-soccer ball literally *does* become a wave when it is not being observed—how else could it interfere with itself? This is a plausible contention, although one that is still subject to debate and interpretation. What seems clear, however, is that each and every nano-soccer ball that makes it past the first barrier somehow does so by going through both slits at once.

Remember, in any given experimental run, there is only one *ball present.*

whether as wave, particle, or wavicle, who can say?

We don't need to take this lying down. In experimental science, when a result is insufficiently clear, the standard remedy is to devise a better experiment—one that tells you more about what is going on. In the double-slit case, why not modify the experiment by monitoring the slits themselves, to observe *which* slit the ball actually goes through? What will we observe now—one ball going through one slit, two balls going through two slits, or some sort of probability wave?

Feel free to hazard a guess yourself, before reading on.

[1] from *The Feynman Lectures on Physics*, vol III (Addison-Wesley, 1965).

The answer was already alluded to on p. 16: we observe only *one ball*, going through only *one slit*. However, there is a price to be paid for our curiosity; the nice, pretty wave interference pattern of Fig. 3.5 is now completely obliterated and replaced with the more mundane particle-like classical probability distribution of Fig. 3.6 (and Fig. 3.4).

We got what we wanted, but we lost what we had.

So, to summarize, not only does the ball go through both slits when we are *not* looking—thereby giving rise to the wave interference pattern—as soon as we *do* observe which slit it goes through, somehow this very act of observation itself destroys the pattern completely. This may seem *crazy*, but in a strange way, it is actually necessary in order to *avoid* paradox.

or perhaps like a bargain with the Devil. . .

In particular, if the interference pattern were to persist even *after* it was clearly established that every ball went through only one slit, what could possibly be the cause? Such a state of affairs would, in effect, constitute a logical contradiction. Instead, quantum particles (it would appear) very kindly spare us from such a catastrophe by putting on their best behavior whenever we are there to watch—in short, by *monitoring us*!

Remember always our very first lesson from p. 14!

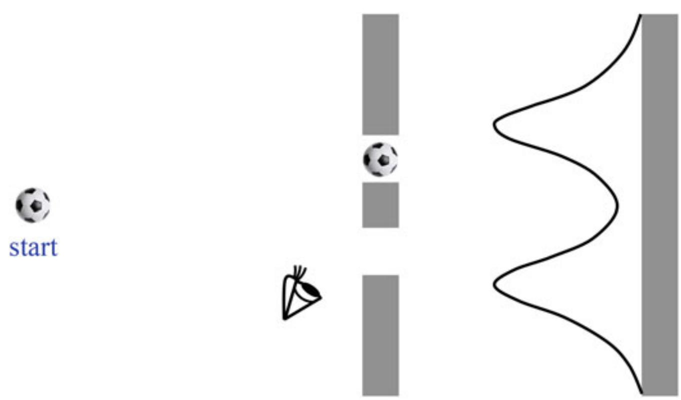


Fig. 3.6 **Quantum double-slit experiment.** The setup is just like Fig. 3.5, except that now we *observe* what happens at the slits.

As a reminder: all of the slit experiments as presented in this chapter may be found on the companion website for this book, in interactive animation form. So, if you did not have at it back on p. 11 (or even if you did. . .), feel free to check these out now.

On the Web: `www.mvjs.org`.

Chapter 4

The Plural Quanta: Expanding the Wavefunction

The double-slit experiment—no matter how much it might "make sense" according to its own weird internal logic—nevertheless may leave you feeling like the victim of a grand conspiracy. Trying to observe exactly how the interference pattern is formed is a bit like trying to see a reflection of yourself with your eyes closed. One may get the distinct impression that quantum particles are mischievous little beings, who only "play nice" when they know they are being watched. If this is how you feel, then hold on to your hat, because things will get a whole lot weirder.

you ain't (not) seen nothin' yet...

First though, we will revisit the wavefunction—since understanding clearly what this entity is, and exactly how it behaves, is evidently essential for a true understanding of quantum mechanics itself. Indeed, much *has* been ascertained over the years—the cumulative benefit of a century's worth of advances in experimental and theoretical methodology. Yet, despite our best efforts to poke and prod it, the true nature of the wavefunction still remains as elusive as ever.

It may not even exist at all.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_4

4.1 Many-Particle Wavefunctions

Let us start with the word *function* and what this word implies.

▷▷▷ **Math Alert!** In mathematics, a *function* is a map, f, that associates a *number* with each element, s, of a set, S. Thus, $s \to f(s)$, where $s \in S$. Usually, the set elements, s, can be uniquely identified using a continuous real-valued variable, x, or a *set* of such variables, $(x, y, z, \ldots)$.

For example, if T denotes temperature, then the function $T(x, y, z)$ represents temperature *as a function of* the position in space (x, y, z). From the description up to this point, one might well imagine that the *quantum wavefunction* (denoted ψ) is a function much like $T(x, y, z)$. Instead of representing the local temperature, though, it presumably represents the probability that a given particle—when measured—will appear at a given position.

This is *kind of* the right idea. Indeed, even scientists often think of ψ in this way, either implicitly or explicitly. Under a bit more scrutiny, however, we find that this idea is quite wrong—in two somewhat subtle, but *extremely* important ways. In my opinion, *most* of the confusion about quantum mechanics—confusion that persists to this day—stems from a failure to keep one or the other of these two points in mind.

The first misconception is simply this:

▷▷▷ **Misconception!!** Quantum mechanics is a theory of individual quantum particles. Each particle, A, B, C, etc., is described by its own wavefunction, $\psi_A(x, y, z)$, $\psi_B(x, y, z)$, $\psi_C(x, y, z)$, etc.

The above idea is nice, simple, intuitive, and just plain wrong. If it were true, then measurement of particles could never be *correlated*. As an example, if I were to measure the position of the proton in a hydrogen atom, then that measurement would tell me *nothing* at all about where its partner electron is located—which, of course, is not the case.

In reality, I can expect to find the electron within a few 10^{-10} m from the proton—with very high probability.

According to statistical theory, in order to incorporate correlations across two different variables, the probability function must be a *joint* function of *both* of those variables.

▷▷▷ **More Accurate !!** The quantum wavefunction describes the entire "quantum system," generally consisting of many particles. *All* particles in the system, A, B, …, are described by a *single* wavefunction, $\psi(x_A, y_A, z_A, x_B, y_B, z_B, \ldots)$.

The upshot is that in quantum physics, you cannot just consider what is happening to a single particle. The wavefunction properly describes only *collections* of particles and is generally *not* reducible into separate pieces describing the individual particles themselves. A quantum particle does not have its own wave.

Conversely, a quantum wave is not strictly associated with just a *single* particle—*nor even with positions in space!!* ψ does not just tell me, say, the probability that particle A will materialize in my left hand, when I measure its position. Rather, it tells me the *joint* probability that A appears in my left hand *and* B appears in my right (assuming the quantum system consists solely of particles A and B).

This is a bit of an oversimplification, but the jist is correct.

Put another way, the "space" (or set, S) on which ψ operates is *not* position space (x, y, z) but is actually far *larger*. It is the space of *all* positions of *all* system particles, $(x_A, y_A, z_A, x_B, y_B, z_B, \ldots)$, known technically as *configuration space.*

Configuration space has 3N dimensions, where N is the total number of particles. You can see why trying to solve quantum problems on a computer quickly becomes a huge challenge…

At this stage, you may well already be asking yourself the most important question that quantum theory must address: just what exactly constitutes "the quantum system"? We will return to this seminal issue later on, but for now, just think of "the system" as a collection of quantum particles, and realize that this collection *as a whole* is what is described by ψ, *not* the individual quantum particles themselves.

and if you really *want to bake your noodle: "does the quantum system include any observers?"*

▷▷▷ **Lesson:** Be very wary of any description of a "wavefunction" that refers to just a *single* particle. This is an *idealization only*, best avoided by the smart quantum consumer.

4.2 Wavefunction Measurement

In Sect. 2.2, we described the wavefunction as a theoretical construct. However, in Chap. 3, we went on to show that ψ can behave like a real tangible wave—e.g., it is capable of interference. Is the wavefunction "real" or not? If so, can it be directly observed? Of course, the answer depends on what is meant by the question. To be precise, I am talking about measuring the wavefunction *itself*—as opposed to measuring the particle(s) that it describes. Also to be precise, by "wavefunction," I am referring to the standard mathematical ψ of standard quantum theory, as described in Sect. 4.1 (and below). With these caveats, the answer to the second question above is a resounding *no!*

though obviously, ψ plays a key role in particle measurement, as discussed...

▷▷▷ **Misconception !!** Using sophisticated new experimental techniques, the quantum wavefunction can now be directly observed in the laboratory—thus definitively proving that the wavefunction really exists.

This is not the case—and if you are getting the hang of this, you realize that it is *not* simply a matter of not yet having the technological capability.

On the Web: `www.nature.com/nature/journal/v474/n7350/full/nature10120.html`.

Wait a minute! How can I be so certain? What about recent, headline-grabbing (and peer-reviewed) scientific papers that claim to have directly measured the wavefunction? Were those experiments done incorrectly? No, the experiments appear to be just fine. It is all a matter of how one chooses to *interpret* the data that was generated—with some interpretations being perhaps a bit more of a stretch than others. One thing, however, is perfectly clear: there is *nothing* in those experiments that suggests any new physics or that otherwise contradicts *any* prediction of standard quantum theory. And yet, standard quantum theory says that ψ *cannot* be directly measured, for reasons I will explain in the latter part of this section.

caveat quantum consumer!

Math Alert! some of this will be a bit technical...

It is important, also, to bear in mind that *no* experiment that has ever been performed has *proven* the existence of ψ. When we hear that something has been "measured," we are naturally inclined to infer that it must therefore surely exist. After all, how could we possibly measure something that is not real? Rest assured, this kind of logic does not apply to wavefunction measurement experiments. To be

ghost hunters also thrive on this sort of logic...

sure, something *is* being measured in those experiments. Moreover, that something *could*, in a sense, be interpreted to be related to a kind of wavefunction—but it could also just as easily be interpreted as something else.

more on this alternative a bit later in this section...

So what *are* these quantum wavefunction measurement experiments actually doing? At the end of the day, these experiments are still extremely impressive and certainly merit a brief discussion here. What they do is to apply a partial measurement of a quantum particle that is *nondestructive*. This means that the wavefunction does not collapse, but it is nevertheless influenced by the measurement—the wave of probability "narrows" down to fewer possibilities but is still a wave. Then, at a later time, a second measurement is performed, causing the wavefunction to collapse. This procedure is then repeated many times, with statistical probabilities gathered for each possible outcome of the second measurement. From these statistics, the influence of the first, weak measurement on the wavefunction can be determined.

Google "quantum tomography" or "weak measurement" to learn more.

This kind of measurement is called a weak measurement*.*

i.e., a conventional, complete or strong *measurement...*

▷▷▷ **More Accurate !!** Using quantum tomography and weak measurement techniques, an "effective" single-particle wavefunction can be mostly inferred, from a large sequence of separate experiments on completely different quantum particles.

Make no mistake, this kind of quantum-level manipulation—i.e., demonstrably "massaging" the wavefunction without causing it to collapse—is very *very* cool. These are great experiments, producing truly spectacular results. As a reality check, though, keep in mind that each individual experimental run is performed on a completely *different* particle. The experiment only "simulates" a single particle (and wavefunction) by attempting to prepare all of the particles in nearly identical initial states. Moreover, all of this assumes that it even makes sense to think in terms of single-particle wavefunctions in the first place—a practice which we know to be somewhat suspect.

The independent existence of $\psi_A(x, y, z)$ would cease the instant particle A became "entangled" with B.

In any event, the broader lesson here is the following, which has not really changed much since Max Born first laid it out for us in the late 1920s:

▷▷▷ **Lesson:** The wavefunction ψ cannot be observed—at least, not directly, not in its entirety, and not as this quantity is defined in standard quantum theory.

presumably noncontroversial *reasons…*

I will now provide some reasons why this is the case, resorting to what standard quantum theory itself says about ψ and its measurement.

The first reason has already been presented—*the* wavefunction describes a *collection* of particles, not individual particles. The second reason is more philosophical: according to the standard quantum theory, the wavefunction is to be interpreted as a theoretical construct that provides information rather than an actual physical entity. It cannot be directly measured because it, itself, is the thing that determines what happens when *any* measurement occurs.

The next reason will first require a bit more explanation about the nature of ψ waves, according to standard quantum theory.

▷▷▷ **Math Alert! *double trouble!!*** The following discussion will be rather technical; some readers may wish to move on to Sect. 4.3. Conversely, others may prefer the *even more* detailed reddit post, listed in the marginal note below.

On the Web: `http://redd.it/1xxmfl`.

In Sect. 4.1, when introducing ψ, I alluded to two subtle and confusing points but only discussed the first one. The second one is this:

▷▷▷ **Misconception !!** The quantum wavefunction ψ is a *probability distribution function*—i.e., a function whose values are positive real numbers everywhere, and whose integration over all space yields one.

▷▷▷ **More Accurate !!** In fact, ψ is a *complex*-valued function, whose *square amplitude* is the probability distribution, i.e., $\rho = \psi^*\psi = |\psi|^2$.

It would be great if we could express all quantum mechanical results in terms of ρ, instead of ψ, but we cannot. Because it maps to *complex* numbers rather than real numbers, ψ contains more information than does ρ—information that turns out to be necessary. Quantum weirdness, it seems, requires the extra information that complex numbers provide.

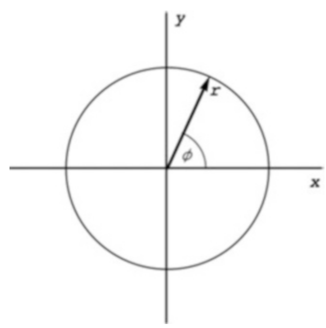
Just what exactly is a complex number, anyway?

▷▷▷ **Math Alert!** You may recall from math class that a single complex number, z, consists of two real numbers, x and y, in the relation, $z = x + iy$, where i is the square root of -1.

Of course, the number -1 *has* no square roots—no *real*-valued square roots, anyway. Thus were "imaginary" numbers, such as i and iy, born (note that iy is the square root of $-y^2$). A *complex* number, then, is one that thus has both real and imaginary parts.

Best cookie fortune ever: "your problems are complex: part real and part imaginary."

One can regard a complex number as a point in the (x, y) plane, as indicated in the marginal figure. For ψ, however, it is better to think in terms of "polar" coordinates (r, ϕ), where r is the modulus or "magnitude" of z and ϕ is the orientation angle or *complex phase*. Note that $r^2 = |\psi|^2 = \rho$ is the probability. The extra, complex phase information, on the other hand—i.e., the "clock" reading on the complex "watch dial"—is a bit harder to interpret physically. However, this is also important, because it is where quantum wave interference comes from.



Real/imaginary (x, y) and magnitude/phase (r, ϕ) decompositions of the complex number, z.

I am indebted to R. Feynman's excellent book for this picture (see footnote).

Specifically, consider two parts of a quantum wave, 1 and 2, that are combined together. If the corresponding complex phase values, ϕ_1 and ϕ_2, are similar, the result is *constructive interference* (i.e., high probability, as in the middle peak in Fig. 3.5). Conversely, if the phases point in opposite directions—2 o' clock and 8 o' clock, say[1]—then the contributions tend to cancel, leading to *destructive* interference (and low probability).

What does this have to do with measuring wavefunctions? Standard quantum theory tells us that only *relative*

[1] R. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton University Press, 1988).

phase—the difference in complex phase angles, $(\phi_2 - \phi_1)$—has physical significance. The *absolute* phase values, ϕ_1 and ϕ_2—i.e., the o' clock values themselves—mean nothing. If I were to "advance the clock" across all parts of ψ by the same amount, then the result would be no change. The new ψ would describe *exactly* the same physical system as the old. Absolute phase is therefore not uniquely defined for a given physical system and can therefore *only* be regarded as a theoretical construct.

This is a simple example of a "gauge freedom," if you have heard that term bandied about.

▷▷▷ **Lesson:** Absolute phase has no physical significance, and can never be measured. At best, all that can be achieved experimentally is an indirect measure of the *relative* phase. Be wary of reports that claim otherwise, or that do not clearly distinguish the two types of phase.

There is one final reason why the wavefunction may not be directly measurable or, at least, why no measurement can prove its existence. That is because it may not exist at all. There are now alternate formulations of quantum theory that use only "trajectories" (paths through configuration space), instead of waves. Moreover, these are *empirically indistinguishable* from the standard wave-based theory—meaning that no experiment can ever determine which theory is correct. Put another way, so long as they continue to validate the predictions of standard quantum theory, then *every* quantum experiment that is performed can be interpreted *either* in wave terms *or* in trajectory terms.

WWW ***On the Web:*** `http://redd.it/1r6kfc`.

although a hypothetical future experiment could *conceivably prove* both *approaches* incorrect. . .

Ironically, this is *even* true for experiments designed to "measure the wavefunction," like those described above! Indeed, similar weak measurement experiments were recently conducted to observe the *trajectory*—rather than wave—aspects of a quantum system. The bottom line is that neither ontology is "proven" or "refuted" by these experiments; they simply offer different vantage points from which one may choose to interpret the same experimental data.

WWW ***On the Web:*** `www.emqm15.org/presentations/speaker-presentations/aephraim-m-steinberg`.

4.3 Collapse and Schrödinger's Cat

We have established that the wavefunction describes a *collection* of particles—being a function of their configuration space, i.e., $\psi = \psi(x_A, y_A, z_A, x_B, y_B, z_B, \ldots)$. What, then, can be said about the individual particles that make up the collection? According to one extreme view, "nothing at all":

▷▷▷ **Misconception !!** (??) A quantum system forms an "undivided whole" whose pieces have no individual existence.

I use question marks here, because there might be a sense in which this could be regarded as a true statement. Many people have used language like "undivided whole" to describe quantum physics—even *bona fide* philosophers and physicists such as D. Bohm.

to say nothing of the more "pop" quantum authors. . .

Nevertheless, in a practical sense, I think it may be more helpful—and certainly less controversial—to state it this way:

▷▷▷ **More Accurate !!** The *context* (environment) of a quantum particle influences its behavior in a way that is totally different from our intuition about the physical world.

Irrespective of the extent to which they can be said to "lose their identity," multiple quantum particles do exhibit some *really* weird behavior, such as *nonlocality*—a kind of long-range interconnectedness. Again, I stress that *nonlocality is an empirical fact*, which has been verified countless times in the laboratory. In any case, the current lesson for the smart quantum consumer is this:

We will explore nonlocality in detail in Chap. 5. . .

▷▷▷ **Lesson:** Quantum physics is all about context.
Note: Though this statement might make some post-modernist intellectuals brim with glee, it has to be—well, taken in the proper context.

So much for particle interconnectedness *inside* the quantum system. What about interactions between the quantum system and the outside world—or more succinctly, between the "observed" and the "observer"? In Sect. 2.2, we saw that these two entities are treated completely differently. One problem is that it is not clear what attributes the observer must possess, in order to justify the special treatment. Can the "observer" be just any other quantum particle? If so, then what makes it different from the system particles? Perhaps the "observer" must be a large or macroscopic object, such as a Geiger counter? Many have postulated that there must be some as-yet-undiscovered collapse mechanism that prevents wave superposition states from persisting up to the macroscopic scale. Others have gone further, suggesting that observers must be conscious beings and/or humans.

i.e., why isn't it simply added to the system?

We will revisit these ideas in Sect. 6.2.

Schrödinger's cat may well be both dead and alive, but his mistresses *were most* definitely *alive. It is said that he developed his famous equation while on a ski trip with one of them.*

A closely related issue is the lack of a natural boundary between the "inside" and the "outside" of a quantum system. This is a serious problem, *because* what happens on either side operates via very different physics. In particular, inside the quantum system, the delocalized ψ evolves deterministically, in accord with the Schrödinger equation (see Appendix I). Outside the quantum system, the observer has the capacity to measure the system, thereby causing a seemingly random collapse of ψ. So, where *does* one draw the line? What happens, for instance, if one places an observer or measurement device *inside* the quantum system? This leads straightaway to Schrödinger's cat.

So, on to the kitties. Schrödinger originally proposed his famous feline thought experiment for the same reason that Einstein proposed EPR (Chap. 5)—i.e., to demonstrate that standard quantum theory leads to absurd results. He did this by simply drawing the quantum "box" large enough to include macroscopic objects. What in effect then has to happen is that macroscopic objects—just like their nanoscale cousins—wind up in an indefinite superposition over macroscopically distinguishable states.

To drive the point home, Schrödinger came up with a rather sadistic but stark example, in which the state of a quantum cat becomes correlated with that of a subatomic particle. As a result, the same superposition of possibilities that describes the particle also describes the cat, which is therefore now in a "wave of superposition" between alive and dead states (Fig. 4.1). Of course, we never observe cat waves nor superpositions of cats that are both alive and dead at the same time. Therefore, according to Schrödinger, standard quantum theory must be wrong or at least incomplete.

You don't know Schrödinger's cat? Really? Google it right now*!*

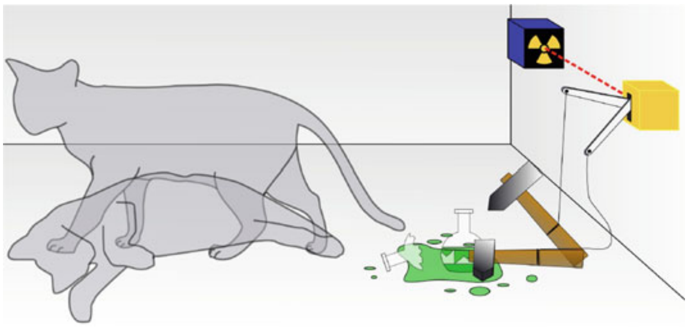


Fig. 4.1 **Schrödinger's cat experiment.** What in blazes is going on? *Quantum superposition*, that's what.

Although the issues raised by Schrödinger do pose a real problem for quantum theory, today we know that his final conclusions were not entirely correct. Like Einstein, Schrödinger believed in hidden variables, and he took it as read that kitties cannot exist in superposition states. The early debate therefore centered on one of two possibilities:

1. The wavefunction must somehow collapse long before the macroscopic scale is ever reached.

2. Quantum theory must be incomplete (e.g., there must be hidden variables that determine a single, definite outcome).

This debate was natural, given the state of affairs at the time. Looking back though, with the benefit of nearly 100 years of hindsight, we now know a few important things that the "founding fathers" did not know (or at least not fully):

It is interesting to speculate how these might have changed the debate at the time, had they been known.

1. Even if a macroscopic observer *were* a part of the quantum system, other objects could still *appear* to that observer to be in definite states.

Google "quantum decoherence" to learn why.

2. Double-slit-type experiments have now confirmed that very "large" objects can be put into superposition states (manifesting as self-interference).

"large" in a quantum sense...

3. EPRB experiments have now confirmed the phenomenon of quantum nonlocality (Chap. 5).

Point 1 above tends to invalidate Schrödinger's original argument, in the following sense: we can no longer rely on our lack of direct observation of macroscopic superposition states as evidence that such states do not persist up to the macroscopic scale.

Point 2 is a *truly* striking development of the last 30 years or so. Naturally, the earliest double-slit experiments were conducted with small things—neutrons, electrons, photons, etc. Researchers then worked their way up to atoms, small molecules, and now, *large* molecules. Indeed, they are now *literally* using nano-soccer balls!! (Chap. 3)—comprised of 60 carbon atoms and *hundreds* of electrons.

in the form of the C_{60} molecule, aka "bucky ball" or "fullerene," which has exactly *the same geometry as a soccer ball.*

It goes way beyond this, even. Very recently, similar experiments were reported using substantially larger biomolecules, comprised of over 800 atoms and *many thousands* of electrons. Moreover, superconducting quantum interference devices* have been used to construct superposition states of *billions* of electrons. Even if one were to regard the most recent experiments as being not yet confirmed, the basic conclusion suggested by these experiments seems clear: it appears there may not be *any* scale beyond which the wavefunction must necessarily collapse. I believe this knowledge would have *astounded* the founding fathers—as much as the third "nail in the coffin" above, i.e., Point 3, the subject of the next chapter.

On the Web: `www.emqm15.org/presentations/speaker-presentations/markus-arndt`.

**affectionately known as* SQUIDs...

If large objects such as cats and people can be superposed, what does this say about parallel universes? See Sect. 6.2.

Separated at Measurement?

Chapter 5

The "Spooky" Quantum: Nonlocality

5.1 Entanglement and Hidden Variables

This chapter deals with *nonlocality* in quantum mechanics—or what is very often called "spooky action at a distance," in deference to a famous Einstein quote.[1] However, Einstein said this while he was still living in Germany, and things spoken in German often seem to have a bit more heft. So this being a particularly weighty topic, I hereby present Einstein's quote in the original German: "spukhafte Fernwirkung."

On the Web: `http://forteana-blog.blogspot.com/2013/05/spooky-action-at-distance.html`.

the German also just sounds plain spookier*...*

As discussed in Sect. 1.2, the quote pertains to the famous EPR thought experiment and "paradox," which Einstein took as proof that standard quantum theory must be incomplete. EPR begins innocently enough, with the creation of two quantum particles that are *entangled*. All that this means is that the combined wavefunction, $\psi(x_A, y_A, z_A, x_B, y_B, z_B)$, cannot be decomposed into separate pieces for the individual particles, $\psi_A(x_A, y_A, z_A)$ and $\psi_B(x_B, y_B, z_B)$. This is the usual situation in quantum mechanics and is therefore not yet controversial; it simply implies that the two particles are statistically correlated.

thereby demonstrating the limits of thought experiments; sometimes you just have to do things in the lab...

The next step is to separate the two particles by a very large distance. A detector then measures particle A, collapsing it down to one definite state. What does this

being careful not to measure them in the process...

[1] Letter from Einstein to Max Born, 3 March 1947.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_5

do to ψ itself? In general, ψ transforms instantaneously into a new wave that describes particle B only—but that depends on the *outcome* of the A measurement. Thus, $\psi \to \psi_B^{(A)}(x_B, y_B, z_B)$. The statistics for a subsequent measurement of B will thus *also* depend on the outcome of the A measurement. This is true *even if* the two measurements are so well separated that not even light can travel between them.

This dependence is indicated by the superscript, "(A)."

Here, then, is the "dilemma," as Einstein saw it. If ψ were the *complete* description of reality, as claimed by standard quantum theory, then *spukhafte Fernwirkung* must be invoked to account for the instantaneous effect on particle B brought about by measurement of particle A—thus violating *relativity theory.* If, on the other hand, ψ merely represents *classical* probability, then instantaneous collapse poses no problem whatever—what has changed is not the state of particle B itself but only our *knowledge* about B. In this scenario, both measurement outcomes are determined in advance and described by hitherto-unknown *hidden variables.* This was the view that Einstein espoused.

i.e., if ψ were a physical entity...

which by that time was completely unchallenged...

This is a bit of an oversimplification, but the jist is correct.

⊳⊳⊳ **Misconception !!** The EPR "paradox" stems from the fact that measurement outcomes of two distant entangled particles, A and B, are *statistically correlated.*

5.2 Bell's Theorem

⊳⊳⊳ **More Accurate !!** Correlation per se is not the relevant issue; the particles could have been correlated right from the beginning. However, this would imply the existence of *hidden variables.*

The above is also known as the "Bertlmann's socks" idea: each sock of a pair is mailed to a different location. One sock package is opened, and the recipient instantly knows the color of the other sock—since pairs of socks always have the same color. Of course, both sock colors were determined all along! Thanks to the sock-color hidden variable, there is no need for one sock to send a faster-than-light signal to the other. This idea is sensible, elegant, and rea-

ironically, though, Bertlmann always wore different *colored socks...*

sonable, and it is no wonder that it had Einstein's seal of approval. However, it also happens to be *completely wrong*, at least in the context of quantum physics.

How do we *know* that it is wrong, since we cannot "see" wavefunctions (or hidden variables) directly? We know this because of something called *Bell's Theorem*. In 1964, John Bell conducted a theoretical analysis of the EPR thought experiment. In particular, he examined statistical correlations in the joint probabilities for combined measurements of A and B. A crucial caveat is that these must include *incommensurate attributes*—such as position and velocity, which are subject to the HUP (Sect. 2.3).

Applying his analysis to local hidden variable theories, Bell discovered that all such theories must satisfy a certain inequality—i.e., $C \leq B$, where C in some sense represents the actual amount of correlation and B is the maximum correlation that classical statistics will allow. By specifying a numerical value for B, Bell's inequality provides an upper bound for C. However, by applying the same analysis to standard quantum theory, Bell discovered that in some cases this predicts... $C > B$! The statistical correlation present in quantum measurements is simply *too great* to be characterized as classical probability.

characterized by classical-like probability waves, of the sort preferred by Einstein...

drum roll, please...

in a sense, this is due to negative probabilities *as we will see in Sect. 5.5*

This was a profound discovery, because it meant that hidden variables were now *experimentally verifiable*. It was not long before experimentalists set out to do exactly that—e.g., J. F. Clauser, A. Aspect, C. Alley, N. Gisin, A. Zeilinger, and others. In what must by now come as no surprise, nonlocal quantum correlation as predicted by standard quantum theory was vindicated—and local hidden variables refuted—in every single case. Such experimental validation of quantum nonlocality has now been pushed to extreme limits. In 1997, Bell inequality violations were observed at a distance of over 10 km. Moreover, if entangled particles are exploiting faster-than-light communication, it has been established that such signals must travel at *at least one million times the speed of light*.

On the Web: `www.emqm15.org/presentations/speaker-presentations/nicolas-gisin/`.

Global *hidden variables may still exist. Recent experiments have focused on closing down a few remaining "loopholes," pertaining to possible measurement device correlation.*

5.3 EPRB as a Game of Coin Toss: Classical Version

although see Appendix I...

A detailed, blow-by-blow account of what happens in the quantum EPRB experiment would be quite involved—requiring a technical exposition of esoteric quantities such as spin that I do not wish to burden you with. At the same time, I do not believe in watering things down any more than is necessary. So, what I have prepared instead is an *exact* description of the EPRB experiment—accurate in every detail, including the numerical values used—except "transposed" into a context that you may find more familiar.

Welcome to the Quantum Coin Toss Game![2]

▷▷▷ **Math Alert!** There will be a tiny bit of mathematics here, but it is nothing more challenging than a Sudoku puzzle. If you have ever solved one of those in the back of an inflight magazine on a short plane ride, then you already have what it takes to become an EPRB master.

I have placed them as far apart on the page as I can, but imagine them to be as distant as you like.

Assume that the detectors never make a mistake, or break down.

We start with the *classical* version of the game. Imagine a pair of identical coin detectors, shown on the left and right sides of Fig. 5.1. There is no connection between them nor any way for them to communicate. The detectors are designed so that every time they accept a coin, they answer a single yes-or-no question about that coin—depending on a user-specified setting. For example, both of the detectors below are currently set to inform us whether the next coin received is "heads" or "tails"—indicated by the upper or the lower light bulb flashing, respectively.

Fig. 5.1 **Classical EPRB experiment.** Two identical coin detectors are both set to register whether the next coin received is "heads" or "tails."

Now imagine a machine that magically conjures up money out of thin air—alas, only nickels and dimes! Also,

[2]With apologies to David Mermin and *Boojums All the Way Through* (Cambridge University Press, 1990).

for some reason it creates Canadian as well as American coins, with either tails up or heads up. Finally, each time you operate the machine, you get a *pair* of coins, which are *identical* in all respects listed above. However, the orientation, nation, and denomination for the coins generated in a given run are completely random. This means that each of the $2 \times 2 \times 2 = 8$ possibilities is equally likely and also that there is no correlation from one run to the next.

two for the price of one... The coins are clearly proxies for entangled quantum particles.

We are ready to begin the game. In Fig. 5.2, we see that the money machine has created two American dimes. The machine tosses one coin in the direction of each detector, which dutifully accepts its respective coin and flashes one light bulb. Based on their current settings, both detectors flash their *upper* bulbs, because both coins are indeed heads up. Of course, we know this because we are now playing *classical* coin toss, which means that we can instantly see all attributes of both coins.

The quantum version of the game is much *more fun.*

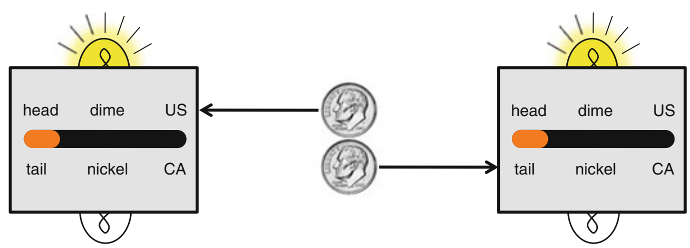


Fig. 5.2 **Classical EPRB experiment.** Two US dimes are created and tossed into their respective detectors, causing both upper bulbs to flash.

Here are a couple more runs, with both detectors still set to measure "tails" or "heads." Both runs happen to be "tails," causing the lower bulbs to flash (Fig. 5.3).

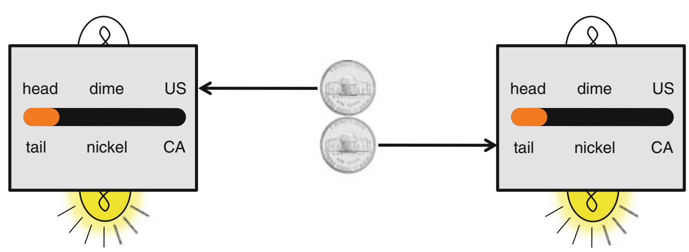


Fig. 5.3 **Classical EPRB experiment.** The next two runs are both "tails."

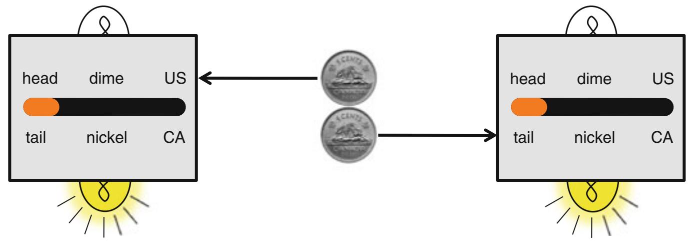


Fig. 5.3 (continued)

Let's make it a bit more interesting, by changing up the settings on our devices—turning them into "nickel/dime" rather than "tail/head" detectors (Fig. 5.4):

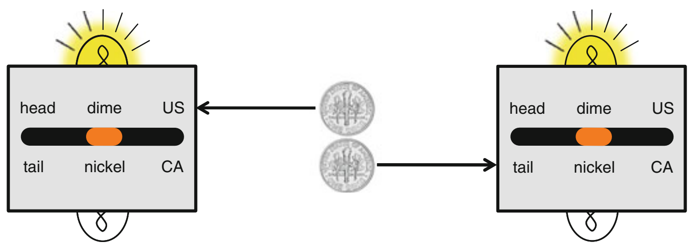


Fig. 5.4 **Classical EPRB experiment.** Both detectors are now set to register whether the next coin received is a nickel or a dime. It is a dime.

You get the idea.

You may have also noticed that for every run so far, the *same* bulb flashed on both detectors. Of course, this is because in every case, both the detector settings and the coins were identical. The only way to get different bulbs to flash in a single run is to use *different* detector settings. So let's play the game one more time, with the left detector on the tail/head setting and the right on nickel/dime. Indeed, this time around, we find that opposite bulbs flash:

Assuming that the two detectors are on different settings, how often can we expect opposite bulbs to flash? Because the runs are completely random, we can easily answer this question by simply making a table of all eight possible coin types. In Fig. 5.6 below, we work through the

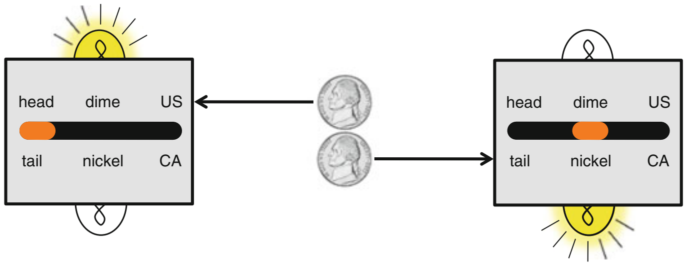


Fig. 5.5 **Classical EPRB experiment.** The only way to get opposite bulbs to flash is to use *different* settings for the two detectors.

case where the detector settings are those of Fig. 5.5. However, the specific attribute settings do not really matter, so long as they are different for the two detectors.

	Coin		Left	Right	Both
head	dime	US	up	up	same
tail	dime	US	down	up	opposite
head	nickel	US	up	down	opposite
tail	nickel	US	down	down	same
head	dime	CA	up	up	same
tail	dime	CA	down	up	opposite
head	nickel	CA	up	down	opposite
tail	nickel	CA	down	down	same

Fig. 5.6 **Classical EPRB experiment.** Outcomes for all eight possible coin types, for detectors set as in Fig. 5.5. Half of these lead to opposite bulbs flashing.

So, the classical EPRB "coin toss" experiment, performed with different detector settings, leads to opposite bulbs flashing exactly 50% of the time. Note that the bulb pattern in the first four lines of Fig. 5.6 is identical to that of the last four lines. This is because neither detector is currently set to discern US vs. CA money, so the behavior has to be the same for both.

If the experiment is repeated many times and 50/50 statistics are not *obtained, the money machine is not completely random.*

5.4 EPRB as a Game of Coin Toss: Quantum Version

Quantum coin toss is *very* different from classical coin toss. To begin with, we cannot "see" the coins directly—as denoted by the blurred gray discs in Fig. 5.7. Our only "eyes" are the detectors themselves. However, we *do* know that the two coins generated in each run of the money machine are still identical. It is therefore prudent to *always* place the detectors on *different* settings. That way, we get *two* pieces of information out of every run, rather than just one. Of course, there is always one attribute that will elude us, no matter how we adjust the settings. On the other hand, to quote a famous celebrity:

"Two Out of Three Ain't Bad."

—*song by Meat Loaf*

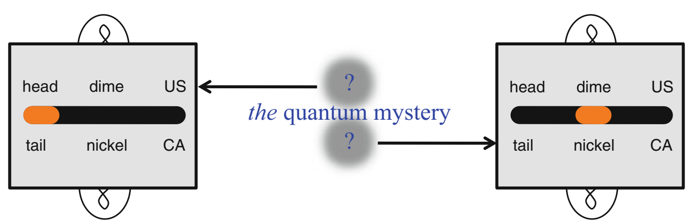


Fig. 5.7 **Quantum EPRB experiment.** Coin attributes can no longer be "seen" directly. Different detector settings now lead to opposite bulbs flashing 75% of the time.

essentially because they are incommensurate *(Sect. 5.2).*

Consider the classical and quantum single-*slit experiments (Chap. 3)... but then again, consider the* double-*slit experiments...*

Although we cannot measure all three attributes for quantum coins, that does not by itself imply that the final statistics will be any different than for classical coin toss. If they *are* the same, that suggests hidden variables—all attributes are well defined at all times, even though we cannot see them. Otherwise, it may suggest *nonlocality* and/or that attributes are determined only at the time of measurement.

according to standard quantum theory...

So which is it in this case? When the settings are different for the two detectors, the quantum statistics are *radically* different from the classical statistics. In particular, opposite bulbs flash *75% of the time*, instead of 50%. This is both *theoretically predicted* and *experimentally validated*. A detailed probability breakdown, based on the only information now available to us (i.e., from the four possible detector outcomes), is provided in Fig. 5.8.

			Detector			Probability	
	Coin		Left	Right	Both	Classical	Quantum
head	dime	?	up	up	same	1/4	1/8
tail	dime	?	down	up	opposite	1/4	3/8
head	nickel	?	up	down	opposite	1/4	3/8
tail	nickel	?	down	down	same	1/4	1/8

Fig. 5.8 **Quantum EPRB experiment.** All four possible measurement outcomes, for detectors set as in Fig. 5.7. Opposite bulbs flash 75% of the time.

The probabilities are now *correlated*; for some reason, quantum mechanics prefers to see nickels with heads and dimes with tails. Why? If you like, this is another manifestation of *the* quantum mystery. But lest we get too far ahead of ourselves, let us now play a new game called "What If?". In this game, we become lawyers or detectives, trying to sniff out every scenario that might account for the unexpected quantum behavior. Perhaps by playing the "What If?" game, we can find some mundane explanation, after all.

or the "WTF" game, if you prefer...

Playing the "What If?" game:

1. Could the quantum results be a statistical fluke? *No.*
2. Could the two quantum coins actually be different? *No.*
3. What if the quantum money maker itself were not completely random? *Leads to negative probabilities.*
4. Could the two detectors be in contact with the money maker? *No.*
5. Could the two detectors (or coins) be in contact with each other? *No.*
6. What about time travel or faster-than-light signaling? *Unlikely.*

5.5 Playing the "What If?" Game

The above is a list of potential "mundane" explanations that have been proposed to account for the experimentally observed quantum EPRB results. In this section we discuss, in turn, the reasons why each of these explanations can be pretty much ruled out.

Explanation 1 is the most obvious: perhaps if we were to conduct many more runs, we might converge to a 50/50 distribution eventually? This suggestion has been thoroughly ruled out—i.e., a sufficient number of experimental runs have been performed to confirm the 75/25 distribution with virtual statistical certainty.

Explanation 2 is also obvious: since we cannot *see* the two coins, perhaps they are not actually the same? We can easily test this idea, simply by going back to using the *same* settings on both detectors. When we do this, we find that the same bulb *always* flashes on both detectors—confirming that the two particles must indeed be identical.

Explanation 3 suggests that the quantum money maker produces coins with *correlated* attributes—which seems inevitable, given Fig. 5.8. The problem is that the probabilities are actually *too* correlated to correspond to *any* classical distribution—which we now demonstrate with the Grand Sudoku Challenge.

or "Magic Cube Challenge" if you prefer...

Assuming all three coin attributes have definite values, then there are eight possible quantum states in all, exactly as in Fig. 5.6. In the classical case, all eight states have the same probability of 1/8. Our goal is to determine the corresponding *quantum* probabilities. It is natural to arrange these into a $2 \times 2 \times 2$ "Rubik's cube," as in Fig. 5.9. Each axis represents a different attribute, and the distance along a given axis represents the attribute value, "down" (0) or "up" (1).

which we know must be different, due to correlation...

For example, the lower left (white) corner of the cube in Fig. 5.9 represents the state "(1,0,0)" or "US, nickel, tails."

Our quantum coin detectors do not give us the eight state probabilities directly, but they *do* provide probabilities for any attribute pair. These are listed for the "denomination-orientation" pair in Fig. 5.8 and also in the lower-left 2×2 square in Fig. 5.9. Detector probabilities for the other two attribute pairs are identical, as seen in the upper and lower-right 2×2 squares in Fig. 5.9. Now, each *detector*

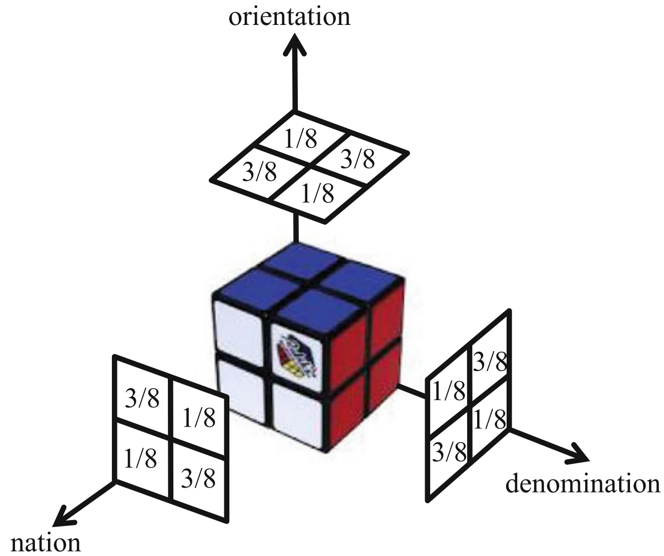


Fig. 5.9 **The Grand Sudoku Challenge.** Find the 8 numbers comprising the $2 \times 2 \times 2$ cube that add up to the numbers indicated in each 2×2 square.

probability must equal the sum of the corresponding *state* probabilities over both values of the unmeasured attribute. For example, $p(0,0,0) + p(1,0,0) = 1/8$, which we know from the last row of Fig. 5.8.

and also from the lower-left corner of the lower-left square of Fig. 5.9...

▷▷▷ **Math Alert!** More generally, *any* pair of adjacent corners in the $2 \times 2 \times 2$ cube of Fig. 5.9 must add up to the value indicated in the corresponding square. The challenge is to find the eight state probabilities that accomplish this. Can *you* solve the Grand Sudoku Challenge?

In fact, there is a *unique* solution. Because of the high symmetry, there are only two distinct types of corners: (1) the "extreme" corners, $(0,0,0)$ and $(1,1,1)$, and (2) all other corners. Thus,

$$\begin{aligned} p(0,0,0) = p(1,1,1) &= a \\ p(\langle \text{all others} \rangle) &= b. \end{aligned}$$

Now we need to only solve for *two* numbers, instead of eight—which is easy to do from the requisite sums:

$$b + b = 3/8 \quad ; \quad b = 3/16$$
$$a + b = 1/8 \quad ; \quad a = -1/16$$

We got what we wanted, but we lost what we had.

Bell's inequality requires the two detector probabilities to be within a ratio of $2:1$*—*exactly *the condition needed to keep both* $a \geq 0$ *and* $b \geq 0$.

So there *is* a solution, but it requires *negative probabilities*, which are not permitted in classical probability theory! Alternatively, the detector probabilities—though all measurable and positive—violate Bell's inequality. These are the "totally insane" results alluded to on p. 6, but in retrospect, they are not so crazy. Once again, as in Sect. 3.5, quantum physics carefully *avoids* paradox—this time, by preventing us from observing states with forbidden probabilities.

What does this have to do with nonlocality? If the correlation is not introduced by the money machine when the coins are created, then it has to come from somewhere else. This might be called "dynamical entanglement," but another term for it might be *cheating*. Like a consistent winner in Las Vegas, this does not happen without some kind of outside communication. In particular, Explanation 4 suggests that the moneymaker and the detectors are "in cahoots"—with the former using skewed statistics to generate coins, based on the settings on the latter.

like loaded dice…

This idea can be mostly ruled out for two reasons: (1) people have checked carefully to make sure that there are no such connections; (2) people have performed *delayed choice experiments*. In (2), the detectors are not set until *long after* the coins have been created and sent on their way. So there is no way for the detectors to communicate their settings to the money machine in advance—even if the devices are somehow mysteriously in contact. Yet even under these circumstances, the same 75/25 statistics are still observed.

Some of these were conducted by C. Alley at the U. of Maryland while I was a student there, though I was not involved.

Another way that the statistics could get skewed is if the two detectors were in communication with *each other*—i.e., Explanation 5. In this scenario, one coin is detected first, and then a signal is sent to the second detector before the second coin is detected. As discussed in Sect. 5.2, however, this scenario is ruled out by experiments in which the detectors are *so* far apart, and the detection times so close, that the purported signal would have to travel orders of magnitude faster than the speed of light.

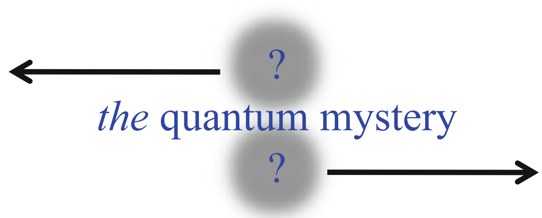
or alternatively, the first coin *sends a signal to the second* coin.

So this is how we come to the last remaining option in our increasingly desperate list—i.e., Explanation 6. No one is quite prepared to abandon relativity theory just yet, and everyone knows time travel can lead to logical paradox. In any case, one thing seems clear: like it or not, *spukhafte Fernwirkung* is here to stay—an uncomfortable fact of life that must be accepted.

although retrocausality, *whereby the future influences the past without paradox, might be a way out...*

As a reminder: all of the EPRB coin toss experiments as presented in this chapter may be found on the companion website for this book, in interactive animation form.

On the Web: `www.mvjs.org`.



Correlated at Birth?

Chapter 6

Where the Weird Things Are

6.1 Where Are They? Where Are We?

Up to this point, I have been careful to emphasize mostly the experimental facts of quantum physics—together with a theoretical structure that is useful for *predicting* those facts. This might be regarded as the "what" of quantum physics—and in this regard, the story that has emerged over the last 100 years is remarkably clear, comprehensive, and unambiguous. Things have reached the point where—for all intents and purposes—we *know* exactly what we will find when we conduct a new quantum experiment.

we are pretty sure*, in any event...*

Yet, these remarkable and highly practical developments tell us very little about the *why* and *how* of quantum physics. As human beings, we have a tendency to want *explanations* for things—causal narratives that progress logically from start to finish, in a way that "makes sense." This, it seems, quantum physics *resolutely* refuses to provide us with—the source of much frustration as well as fascination.

As Feynman puts it, we have only a description *rather than an* explanation.

in short, stories...

Consider that the basic situation, as it has unfolded up till now, is essentially this:

1. Quantum physics places limits on what we can directly observe.

This manifests in both EPRB and double slit—though in different ways.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*, Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_6

Talk about "what you don't know can't hurt you"!

2. If we *could* see what we are prevented from seeing, the result would be logical contradiction and/or paradox.

one with a sense of humor, if without mercy...

What kind of a perverse world is this? What Creator would conjure up such a thing? The more we are told to pay no attention to the man behind the curtain, the more we *have* to look. Since the Magician will never allow us to sneak a peek backstage, however, all we can do is content ourselves with "conspiracy theories."

Such conspiracy theories about quantum mechanics are more typically referred to as *interpretations*.

▷▷▷ **Danger! Will Robinson, Danger!!**
We are now entering the realm of *METAPHYSICS!!*

This, of course, is what causes all the real trouble...

Metaphysics is the branch of philosophy that deals with the fundamental nature of What Is. Its origins trace back to ancient Greece—and certainly, Point 1 above is reminiscent of Plato's famous Allegory of the Cave. Point 2, on the other hand, is much more akin to Pandora's Box! The point is simply that if we seek a deeper understanding of the quantum mystery, we are pretty much forced into the realm of philosophy, rather than science. Quantum physics very naturally leads us *up* to that realm but provides little in the way of a roadmap to guide us, once we have entered.

Of course, there are many who feel that we should simply *not* enter.

"What cannot be seen should not be discussed."

—*philosophy of Niels Bohr*[1]

"Abandon hope all ye who enter here."

—*Dante's* Inferno

[1] From Fred Alan Wolf, *Parallel Universes* (Simon and Schuster, 1988)

6.2 Interpretations of Quantum Physics

Niels Bohr, of the first quote above, was one of the chief architects of the *Copenhagen Interpretation*. This is the *standard* interpretation. It is not so much an explanation of the deeper meaning as it is a stance that such a thing should not be pursued. Some scientists prefer Copenhagen because they believe it comes with the least "metaphysical baggage." Others find it to be a "cop-out" that merely sweeps this baggage under the rug. Regardless, Copenhagen does make certain metaphysical claims. One of these is that wavefunction collapse occurs as a result of measurement. Another is that physical systems do not actually *possess* attributes until those attributes are measured.

associated with what I have referred to throughout this narrative as the "standard quantum theory"...

or seems *to, at any rate; there is a bit of ambiguity about it all...*

no hidden variables...

Dissatisfied with Copenhagen's deliberately limited perspective, many efforts to adopt a deeper view have been undertaken over the years. A comprehensive description—even of just the major interpretations—would lie far beyond the scope of this narrative. Instead, I will attempt to provide some flavor of how they work by categorizing them based on a simple question: where does wavefunction collapse occur?

Google, e.g., "de Broglie-Bohm," "many worlds," "transactional interpretation," "stochastic mechanics," "ensemble interpretation," among others.

An important point is that all interpretations for the most part correspond to the same experimental predictions—i.e., those of the standard quantum theory. They differ only in terms of their explanation of what is happening *behind* the scenes. These differences are therefore metaphysical, and judged by qualities other than agreement with experiment—simplicity, utility, aesthetics, etc. In that respect, the situation can be likened to the geocentric vs. heliocentric (Copernican) models of the solar system.[2]

i.e., they offer different possible "tricks" to explain the same magic act...

These two models can *exhibit subtle empirical differences! This was even known at the time and used against Copernicus (see footnote). Moreover, the same is true for some quantum interpretations.*

Let us return to our simple question: where does wavefunction collapse occur? There appear to be three primary options, none completely satisfactory:

1. Collapse occurs randomly at some small threshold scale, due to new physics that has yet to be discovered.

[2]D. Sobel, *A More Perfect Heaven: How Copernicus Revolutionized the Cosmos* (Walker & Company, New York, 2011)

2. Collapse occurs with the first conscious observer.

3. Collapse never occurs. Everything, including observers, is a part of the quantum system.

Option 1 describes the so-called "objective collapse" theories—e.g., of Ghirardi, Rimini, and Weber (GRW), and R. Penrose. From a theoretical standpoint, this seems the most reasonable and the least conceptually difficult. It has the advantage that it does not require a subjective observer nor does it allow macroscopic superposition states. The problem is that the "threshold" seems to keep getting higher and higher, as experiments self-interfere larger and larger objects. We have already reached biomolecules and billions of electrons (Sect. 4.3). At what point do we admit that we have effectively reached the size of cats?

Googling "quantum" and "conscious" together yields 775,000 hits.

The fate of Schrödinger's cat hangs in the balance.

Don't expect this to be much the same concept as in mysticism or Eastern religions, however.

Option 2 has spun off a whole "new age" cottage industry (see Sects. 1.3 and 6.3). It also has a surprising number of adherents even among scientists one would not necessarily expect, such as J. von Neumann. However, the obvious major flaw from a scientific standpoint is that it is not clear what a "conscious observer" is. Does this mean "human" for example? All of that said, consciousness research, as a legitimate subfield of psychology and cognitive science, is growing rapidly these days. Such work may well shed light on quantum measurement or at least provide a more scientific definition of consciousness.

In the 1950s, H. Everett III, the inventor of this idea, called it the "universal wavefunction."

which, again, the interested reader can Google...

This term "many worlds" was popularized not by Everett, but later by B. DeWitt, in the 1970s. It was a different time.

Finally, there is option 3. From a purely a theoretical standpoint, this approach has many advantages. In particular, there is no need to introduce new physics nor to worry about where to draw the boundary between quantum system and observer. The "**When the particle is *not* being observed**" paragraph of Sect. 2.2, *by itself*, provides the complete description. The "**When the particle *is* being observed**" paragraph can be *almost* ignored—only "almost" because the *illusion* or perception of wavefunction collapse must be addressed. However, this can be very satisfactorily explained through *decoherence*.

On the other hand, there is the slight hiccough that absolutely *everything* becomes a part of the wave superposition—meaning that there are many copies of everything out there, including people, including you and me. Thus was the so-called *many-worlds interpretation* born.

Such ideas have greatly captured the public fascination, particularly recently. However, many scientists find the approach distasteful. They regard the proliferation of worlds to be "extravagant"—or perhaps they are just not comfortable with the idea that they themselves might have *doppelgängers.*

Googling "many-worlds theory" yields 1.54M hits.

who may even disagree with their own views on many worlds!

6.3 Popular Depictions of Quantum Physics

More than for any other quantum interpretation, critics like to point out that the many-worlds interpretation can never be validated experimentally. While this is likely correct, as we saw in Sect. 6.2, the same criticism could be leveled against *most* interpretations.

⊳⊳⊳ **More Accurate !!** The different interpretations of quantum physics are (mostly) characterized by the same empirical predictions and can therefore *never* be "proven" by experiment.

Because of this simple fact, the debate about *the* "right" or "best" interpretation will doubtless continue for many years to come. In my view, they (mostly) all deserve their day in the sun, as they each offer unique and valuable insight. Moreover, it is perfectly fine to speculate about what is happening behind the scenes—so long as it is clear that this is what one is doing and, also, that said speculation agrees with experiment.

The interpretive "trick" must properly account for the observed "magic act."

Yet, things *often* go astray in popular treatments with the following:

⊳⊳⊳ **Misconception !!** Quantum experiments have now "proven" the existence of entity x from interpretation y.

We have grown accustomed to—perhaps even jaded by—headlines of the "Scientists prove existence of parallel universe" variety. Don't buy it. Unless x is a universal entity, common across *all* interpretations (and theoretical formu-

remember to be a smart quantum consumer…

lations), its existence is decidedly *not* proven by quantum experiments (see, e.g., Sect. 4.2).

Even worse is the following:

▷▷▷ **Misconception !!** The different interpretations of quantum mechanics are all true at the same time.

▷▷▷ **More Accurate !!** Each interpretation provides a consistent framework on its own, but they cannot be "mixed and matched."

You might as well put a Volvo cylinder head on a Ford engine block.

One often sees a mishmash of elements from different interpretations, jumbled together as if they were all true at once—many worlds combined with conscious collapse, for instance. One cannot just arbitrarily combine bits and pieces from different interpretations, without losing internal consistency. That's not "quantum weirdness," it's just plain wrong. Unfortunately, it is also quite common in some popular depictions.

Some really do inform; others just sensationalize.

I could discuss the merits and weaknesses of popular depictions until I am blue in the face, but that would not address what I regard to be the bigger issue: why are they so popular in the first place? What causes the enormous demand discussed in Sect. 1.3? The strong sense that I get—from internet comments, interviews, and one-on-one meetings—is simply this: *People have a need for science to provide meaning and/or to legitimize their experiences and world views.*

for some reason, Hamlet *comes to mind. . .*

Is such a need misplaced? One message that comes through loudly in many popular depictions is that *there is more out there than just the ordinary mundane reality.* I think this is important and healthy: none of us need be slaves to the sort of rote behavior and automatic thinking to which we are constantly exposed. But that is being a smart *consumer*, period—we don't have to drag quantum physics into it. Put another way, quantum physics may serve as an excellent *metaphor* for various other things that give us meaning—spirituality, transcendent religious experience, sense of nonlocal interconnectedness, or connection with a deity—without having to be regarded as their *causative agent* or as *proof* that they exist.

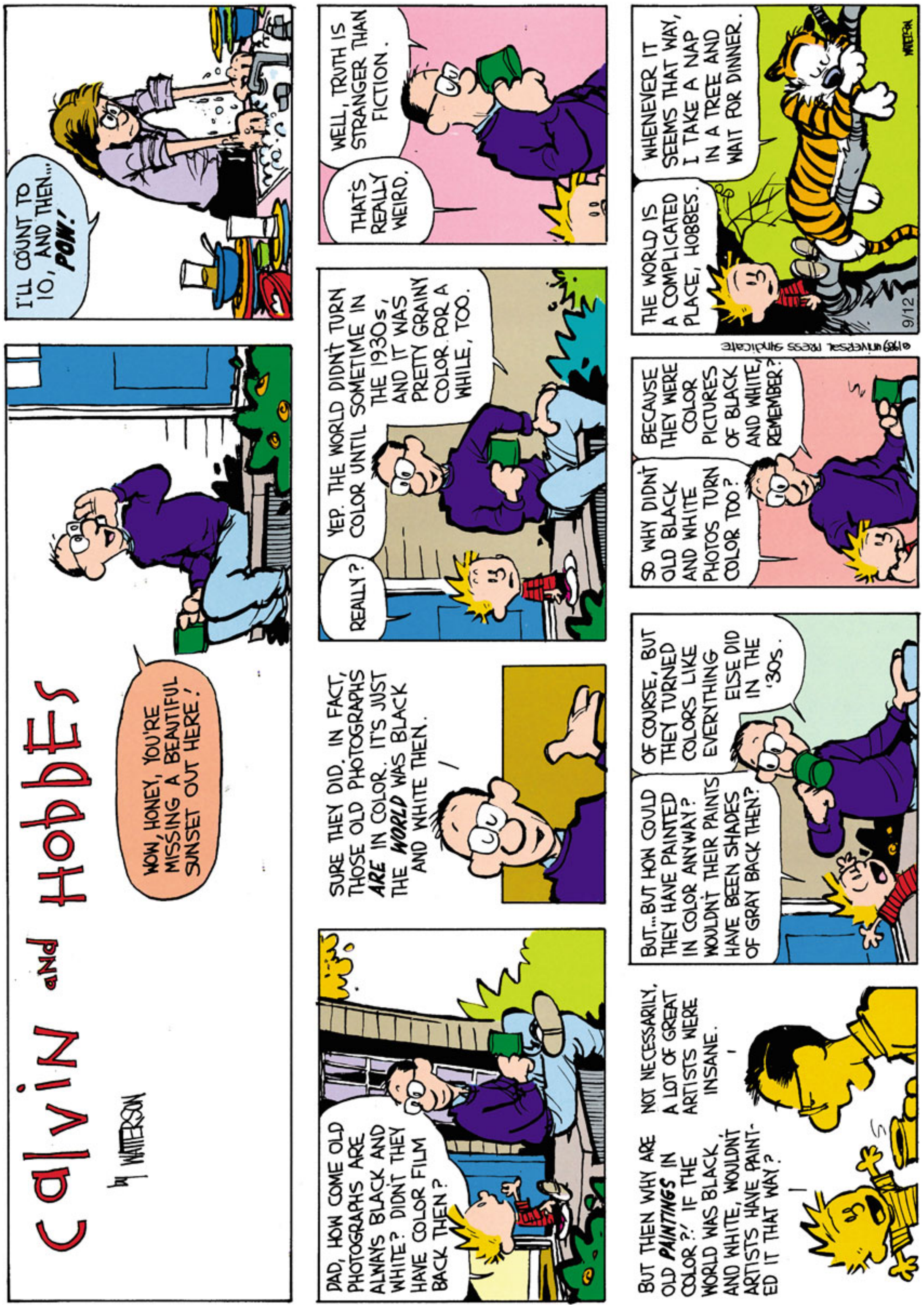
In my view, the legitimacy of such experiences requires *no such proof.*

In conclusion, scientists do not *own* science; it is part of the great public domain of human knowledge that we all share. Moreover, science is *interesting*, and enormously important in our lives. So it is right that people should be fascinated by it and want to learn about it. By offering nonexperts the tools needed to form their own legitimate opinions, popular depictions can provide a great service—but only if they take the necessary pains to "get it right." *Caveat quantum consumer!!*

a Top 10 List of Most Cringeworthy Quotes is available on request...

▷▷▷ **Lesson:** Speculation is fine, *but*... it should not hide behind the mantle of "hard scientific fact."

▷▷▷ **Lesson:** Quantum physics is not just for physicists, *but*... the accepted reality is mysterious and wonderful enough on its own, without the need for embellishing hype.


Calvin and Hobbes
by WATTERSON
WOW, HONEY, YOU'RE MISSING A BEAUTIFUL SUNSET OUT HERE!
I'LL COUNT TO 10, AND THEN... POW!
DAD, HOW COME OLD PHOTOGRAPHS ARE ALWAYS BLACK AND WHITE? DIDN'T THEY HAVE COLOR FILM BACK THEN?
SURE THEY DID. IN FACT, THOSE OLD PHOTOGRAPHS ARE IN COLOR. IT'S JUST THE WORLD WAS BLACK AND WHITE THEN.
REALLY?
YEP. THE WORLD DIDN'T TURN COLOR UNTIL SOMETIME IN THE 1930s, AND IT WAS PRETTY GRAINY COLOR FOR A WHILE, TOO.
THAT'S REALLY WEIRD.
WELL, TRUTH IS STRANGER THAN FICTION.
BUT THEN WHY ARE OLD PAINTINGS IN COLOR?! IF THE WORLD WAS BLACK AND WHITE, WOULDN'T ARTISTS HAVE PAINTED IT THAT WAY?
NOT NECESSARILY. A LOT OF GREAT ARTISTS WERE INSANE.
BUT... BUT HOW COULD THEY HAVE PAINTED IN COLOR ANYWAY? WOULDN'T THEIR PAINTS HAVE BEEN SHADES OF GRAY BACK THEN?
OF COURSE, BUT THEY TURNED COLORS LIKE EVERYTHING ELSE DID IN THE '30s.
SO WHY DIDN'T OLD BLACK AND WHITE PHOTOS TURN COLOR TOO?
BECAUSE THEY WERE COLOR PICTURES OF BLACK AND WHITE, REMEMBER?
©1989 Universal Press Syndicate
THE WORLD IS A COMPLICATED PLACE, HOBBES.
WHENEVER IT SEEMS THAT WAY, I TAKE A NAP IN A TREE AND WAIT FOR DINNER.
9/12
WATTERSON

Appendix I: Looking Under the Hood: The Mathematical "Method" Behind the "Madness"

▷▷▷ **Math Alert!** ***double trouble!!*** In this Appendix, we present some of the mathematical equations that form the bedrock of what we have many times referred to as the "standard quantum theory." Warning! Not for the faint of heart.

The most important equation in all of quantum mechanics is also the simplest:

$$\hbar = 1.0545718 \times 10^{-34} \text{ J s} \tag{6.1}$$

However, like all equations, it requires an explanation for a proper understanding. The quantity $\hbar$ is called (the reduced) *Planck's constant*. It is a *fundamental* constant of nature, representing the "size" of the quantum. In Eq. (6.1), $\hbar$ is given in standard macroscopic units called "SI" (*système international*) units—in terms of which, it is seen to be very small indeed. The *dimensions* of Eq. (6.1) are also significant; we see that $\hbar$ has units of energy × time, also known as *action*.

More generally, pairs of quantities whose product has dimensions of action are called *conjugate variables*. Other examples include position × momentum and angle × angular momentum. Conjugate variables are important because they are *incommensurate* and can therefore be used to formulate uncertainty principles. Thus, the usual HUP as described in Sect. 2.3 is actually the position–momentum uncertainty principle:

$$\Delta x \Delta p \geq \hbar/2, \tag{6.2}$$

where Δx is the uncertainty (technically *standard deviation*) in position, and Δp is the uncertainty in momentum (where momentum is just mass × velocity). The energy–time and angle–angular momentum uncertainty relations have a similar form.

From a wave mathematics point of view, the uncertainty principle is nothing new. Given any "wave" function $\psi(x)$, the *Fourier transform* function, $\tilde{\psi}(k)$, is defined:

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-ikx)\psi(x)\, dx \tag{6.3}$$

This represents the decomposition of the original wave into its Fourier (i.e., sinusoidal) components. It is well known that a narrow $\psi(x)$ implies a broad $\tilde{\psi}(k)$ and vice versa. In particular,

$$\Delta x \Delta k \geq 1/2. \tag{6.4}$$

Equation (6.4) suggests the identification of p with $\hbar k$, which has the right dimensions for momentum (since k must have dimensions of $1/x$, and x has dimensions of length). Note that $\lambda = 2\pi/k$ is the *wavelength* for the Fourier component k. Thus,

$$p = \hbar k = 2\pi\hbar/\lambda; \tag{6.5}$$

momentum in quantum mechanics is inversely proportional to wavelength.

Using wave properties of ψ as discussed above, we can estimate the distance f between adjacent "fringes" in the double-slit interference pattern that appears on the far wall in Fig. 3.5. First note that the *phase* of the ψ branch that passes through the upper slit is given by $\phi_+ = kx_+$, where x_+ is the distance from the upper slit. Likewise, $\phi_- = kx_-$ for the lower slit. Since *relative* phase is what causes interference (Sect. 4.2), f must correspond to a relative phase change of one cycle, i.e.,

$$\Delta(\phi_+ - \phi_-) = k\Delta(x_+ - x_-) = 2\pi. \tag{6.6}$$

If s is the distance between the two slits, and D the distance between the two walls, with $s/D \ll 1$, then simple geometry shows $\Delta(x_+ - x_-) \approx f(s/D)$, or

$$f \approx 2\pi D/sk = \lambda(D/s). \tag{6.7}$$

The above "kinematic" description is useful for interpreting $\psi(x,t)$ once we have it but does not tell us how $\psi(x,t)$ actually changes over space and time. A complete theory requires such a *dynamical* rule—which in this case should take the form of a *partial differential equation* (PDE). Fourier analysis can be used to guess the correct PDE. In particular, k is associated with the *partial derivative* of its *conjugate variable*:

$$k \to -i(\partial/\partial x), \quad \text{or} \quad p = \hbar k \to -i\hbar(\partial/\partial x). \tag{6.8}$$

Likewise, energy, E, can be associated with *its* conjugate variable, i.e., $E \to i\hbar(\partial/\partial t)$. Substituting these relations into the classical "kinetic-plus-potential" energy expression

$$E = \frac{p^2}{2m} + V(x), \tag{6.9}$$

and applying the result to $\psi(x,t)$, we obtain the desired quantum PDE:

$$i\hbar\left(\frac{\partial\psi}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2}\right) + V(x)\psi. \tag{6.10}$$

Equation (6.10) is the celebrated (time-dependent) *Schrödinger equation*, in 1D (one spatial dimension). Note the presence of $i = \sqrt{-1}$ on the left hand side, which ensures that $\psi(x,t)$ is necessarily complex valued. In 3D space, Eq. (6.10) generalizes to

$$i\hbar\left(\frac{\partial\psi}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + V\psi = -\frac{\hbar^2}{2m}\left(\nabla^2\psi\right) + V\psi. \tag{6.11}$$

Here, the potential $V = V(x,y,z)$, and $\psi = \psi(x,y,z,t)$. Of course, this describes only a *single*-particle wavefunction. The *real*, i.e., many-particle Schrödinger equation is

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m_A}\nabla_A^2\psi - \frac{\hbar^2}{2m_B}\nabla_B^2\psi + \cdots + V\psi, \tag{6.12}$$

with $\psi = \psi(x_A, y_A, z_A, x_B, y_B, z_B, \ldots, t)$ and $V = V(x_A, y_A, z_A, x_B, y_B, z_B, \ldots, t)$.

In addition to spatial coordinates (x,y,z), quantum particles also possess intrinsic *spin* attributes. For any direction in space—e.g., such as the vertical z axis, $(0,0,1)$—a spin measurement yields (for most particles) one of two possible outcomes: "up" or "down." Spin can be measured in any direction; however, different spin measurements are *incommensurate*. Note that this situation corresponds closely to the quantum coin toss experiment of Sect. 5.4. One minor technical difference is that the two entangled quantum particles in an EPRB experiment are prepared with *opposite*, rather than identical, spin values. The three detector settings correspond to three different spin directions, 120° apart in a single plane. It can be shown that if the detector settings are different, and the first detector registers "spin up" for particle A, then ψ collapses partially (as in Sect. 5.1) to the following superposition over B states only:

$$\psi_B^{(A=\mathrm{up})} = \left(\frac{1}{2}\right)\psi_{\mathrm{up}} + \left(i\frac{\sqrt{3}}{2}\right)\psi_{\mathrm{down}} \tag{6.13}$$

Since probabilities are obtained as square amplitudes, the probability that subsequent measurement of B will find it to be in a spin-up state is $p_B(\mathrm{up}) = |1/2|^2 = 1/4$. Likewise, $p_B(\mathrm{down}) = |i\sqrt{3}/2|^2 = 3/4$. Note that these are the probabilities for the *same* and *opposite* bulbs flashing on the two detectors, respectively.

Appendix II: Further Reading

Less Advanced Books (alphabetical order by author)

I. Asimov, *The New Intelligent Man's Guide to Science* (Basic Books, New York, 1965)

J.E. Baggott, *Farewell to Reality: How Modern Physics Has Betrayed the Search for Scientific Truth* (Pegasus Books, New York, 2014)

A. Einstein, *The World As I See It* (Citadel Press, 2000)

R.P. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton University Press, Princeton, 1988)

R.P. Feynman, *Six Easy Pieces* (Perseus Books, Cambridge MA, 1995)

R.P. Feynman, *Six Not-So-Easy Pieces* (Perseus Books, Cambridge MA, 1997)

R.P. Feynman, *The Character of Physical Law* (British Broadcasting Corp., 1965)

B. Greene, *The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos* (Alfred A. Knopf, New York, 2011)

G. Greenstein, A.G. Zajonc, *The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics* (Jones and Bartlett, Sudbury, 1997)

M. Kafatos, R. Nadeau, *The Conscious Universe* (Springer, 1990)

D. Mermin, *Boojums All the Way Through* (Cambridge University Press, Cambridge, 1990)

A. Pais, *Inward Bound: Of Matter and Forces in the Physical World* (Clarendon Press, Oxford, 1988)

R.A. Bertlmann, A. Zeilinger (eds.), *Quantum [Un]speakables: From Bell to Quantum Information* (Springer, 2002)

L. Smolin, *The Trouble with Physics* (Mariner Books, 2007)

D. Sobel, *A More Perfect Heaven: How Copernicus Revolutionized the Cosmos* (Walker & Company, New York, 2011)

S. Weinberg, *The Discovery of Subatomic Particles* (Cambridge University Press, Cambridge, 2003)

F.A. Wolf, *Parallel Universes* (Simon and Schuster, New York, 1988)

More Advanced Books (alphabetical order by author)

J.E. Baggott, *The Meaning of Quantum Theory: A Guide for Students of Chemistry and Physics* (Oxford University Press, 1992)

J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1988)

D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, 1951)

E. Commins, *Quantum Mechanics: An Experimentalist's Approach* (Cambridge University Press, New York, 2014)

P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, 1958)

R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics*, Vols. I–III (Addison-Wesley, 1965)

M.S. Longair, *Theoretical Concepts in Physics* (Cambridge University Press, Cambridge, 1984)

L. Pauling, E.B. Wilson, *Introduction to Quantum Mechanics With Applications to Chemistry* (Dover, New York, 1963)

J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1983)

H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950)

Internet Sites (in order referenced, plus additional)

http://www.amazon.com

http://www.mvjs.org

http://www.google.com

http://www.emqm15.org/presentations/speaker-presentations/

http://www.nature.com/nature/journal/v474/n7350/full/nature10120.html

http://redd.it/1xxmfl

http://redd.it/1r6kfc

http://forteana-blog.blogspot.com/2013/05/spooky-action-at-distance.html

http://www.acoustics.salford.ac.uk/feschools/waves/diffract3.php

http://www.huffingtonpost.com/bill-poirier/quantum-weirdness-and-many-interacting-worlds_b_6143042.html

http://physics.stackexchange.com

Appendix III: Glossary

Bell's Theorem: Theorem derived by John S. Bell in 1964 establishing the experimental decidability between quantum theory and local hidden variable theories; quantitatively it is expressed as an inequality known as "Bell's inequality."

Classical Mechanics: Theory of physics epitomized by the work of Sir Isaac Newton in the seventeenth century that describes a "clockwork" motion of objects; characterized by precision and "commonsense" determinacy; later superceded (at very small scales) by quantum mechanics.

Double-Slit Experiment: Experiment demonstrating the wavelike nature of quantum particles, by firing a large number of identically prepared particles, one at a time, at a screen with two separated parallel slits, and observing a pattern of interference fringes on a second screen placed beyond the first.

Einstein Podolsky Rosen [Bell] (EPR[B]) Experiment: Thought experiment originally conceived by Albert Einstein and coworkers in 1935 (EPR), to demonstrate flaws in quantum theory, and prove the existence of hidden variables; later, in the wake of Bell's Theorem, EPRB became a *bona fide* laboratory experiment—which ruled out local hidden variables and vindicated quantum nonlocality.

Entanglement: In quantum mechanics, refers to two or more particles that are statistically correlated, even across vast distances, such that they cannot be described by separate single-particle wavefunctions.

Heisenberg Uncertainty Principle (HUP): Principle formulated by German physicist Werner Heisenberg in 1927 that describes the indeterminacy of quantum mechanics; qualitatively, it states that there is a fundamental limit to the precision with which both the position and velocity of a quantum particle can be simultaneously determined; quantitatively it can be expressed in various mathematical equations, generally as an inequality.

Hidden Variables: Dynamical attributes purported by some to be missing from quantum theory, which would render it complete and deterministic; *local* hidden variables of the sort advocated by Einstein are now ruled out by EPRB experiments, although *global* (nonlocal) hidden variables may still exist.

Misconception: A mistaken view; an erroneous notion; impossible for human beings to avoid completely.

Nonlocality: Principle of physical theories that postulates "action at a distance," i.e., the remote influence of one particle on another, either instantaneously, or at speeds faster than light; in quantum mechanics, this concept can be a bit more subtle (see Entanglement).

Particle: Refers to "localized" (point-like) objects in the physical world; in classical physics, particles trace out a definite "trajectory" (one-dimensional curve through space) over time; in quantum physics, a "particle" can manifest delocalized wavelike behavior, when its position in space is not being observed.

Probability Wave: Mathematical function describing the likelihood that a particle is located at a given point in space (may also refer to sets of particles); used in both classical and quantum mechanics, though the behavior is quite different in each case.

Quantum Mechanics: Theory of physics originating in the early twentieth century that describes the mechanics of atoms, molecules, etc.; replaces the earlier "classical" theory of mechanics; characterized by indeterminacy; impossible for human beings to understand completely.

Schrödinger's Cat: Thought experiment devised by Austrian physicist Erwin Schrödinger in 1935, to demonstrate the absurdities that arise from quantum theory when the notion of a superposition state is extended up to the macroscopic realm.

Wave: Any entity that is delocalized over a broad region of space.

Wavefunction: Complex-valued mathematical function that describes the state of a quantum system; the square amplitude of the wavefunction is the probability wave.

Wavefunction Collapse: Process by which measurement of a quantum system by an outside observer causes the probability wave to "collapse" to one specific state from among a superposition of states—seemingly at random.

Wave-Particle Duality: Refers to the notion that quantum objects may exhibit different behaviors, under different circumstances, reminiscent of either classical waves or classical particles.

Credits

Photos:

A. Einstein: Arthur Sasse / AFP-Getty Images / Wikimedia
R. Feynman: The Nobel Foundation / Wikimedia / Public Domain
Y. Berra: Baseball Digest / Wikimedia Commons / Public Domain
W. Heisenberg: Friedrich Hund / Wikimedia Commons / CC-BY-3.0
W. White, aka "Heisenberg": "Breaking Bad" / Vince Gilligan / AMC / Wikimedia
E. Schrödinger: The Nobel Foundation / Wikimedia / Public Domain
Prudy (cat): Photos by the author.

Other Images:

skiier: Quantum Skiier / Courtesy of Anne Longo / derived from a cartoon by Charles Addams / The New York Times (December 3, 2006)
Yogi Bear: Hanna-Barbera Productions / Wikimedia
cat: Diagram of Schrödinger's Cat / D. Hatfield / Wikimedia / CC-by-SA 3.0

Quotes: All quotations are cited where used.

Other:

"Separated at Birth?": Former Reg'd Trademark / Spy Publishing Partners
cartoon: Bill Watterson / Calvin and Hobbes © (October 29, 1989) / Reprinted with permission of UNIVERSAL UCLICK / All rights reserved

Part II: Misconceptions About Particles and Spacetime

John Terning

Introduction

I'm going to start out Part II of the book with a review of some basic misconceptions: many people think of elementary particles as tiny little balls, they envisage atoms that look like tiny solar systems, and they also imagine that to see atoms they would need a very large microscope. After covering these basics, I'm going to move on to some topics that have been ripped from news headlines, such as "Collider Spawns Planet-Devouring Black Hole," "God Particle Could Wipe Out The Universe," "LHC Could Prove Existence of Star Trek's Parallel Universe," "Scientists Raise Concerns About Cell Phones," and "The Cold Fusion Race Just Heated Up." I'll try to examine the scientific ideas behind these hyperbolic headlines and give some more depth so that the reader can understand the real issues involved. For the mathematically comfortable reader, an extra chapter is included that discusses these topics using high school math.

Chapter 7

Particles

7.1 What Are Particles?

The reason that a lot of people think of elementary particles as being tiny little balls is that scientists often show pictures of atoms and elementary particles where the particles are represented by tiny little balls, see Fig. 7.1. As far as we know, elementary particles are fluctuations of quantum fields, but that is very hard to convey in a picture.

The ancient Greeks conceived of atoms as elementary particles that cannot be sub-divided. Thousands of years later, people were able to probe deeper into the substructure of nature and found that atoms are actually made up of a very small nucleus surrounded by cloud of electrons. It was later found that the nuclei of atoms were made of even smaller particles: protons and neutrons. Eventually we found out that protons and neutrons themselves were made of even more elementary particles called quarks.

Democritus (c. 460 BCE-c. 370 BCE) said: "but in truth there are only atoms and the void."

Ernest Rutherford (1871–1937) discovered the nucleus in 1911.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_7

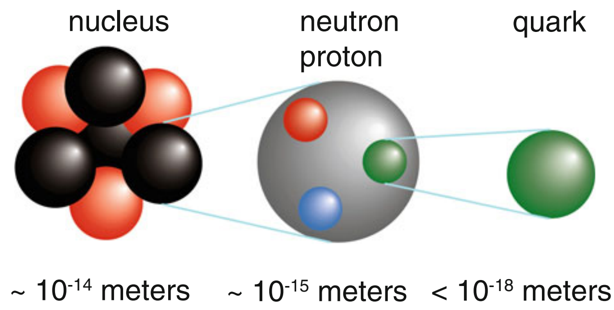


Fig. 7.1 The core of the atom is a nucleus, the nucleus is made of protons and neutrons, which in turn are made of quarks. While experiments have be able to resolve the size of the nucleus, the proton, and the neutron, no one has seen any evidence that quarks (and electrons) have a size; we only know their size is less than 10^{-18} meters. The picture is obviously *not* drawn to scale.

▷▷▷ **Misconception !!** Particles are tiny little balls.

As far as we know quarks (and electrons) are elementary but as with all scientific knowledge we don't know that for certain. Our knowledge always precedes as a series of successive approximations. The experiments we've done so far don't show any evidence for quarks having a substructure, but at different times in the past that same statement was true of atoms, nuclei, and protons.

A meter is about the same length as a yard. 10^{-18} *is shorthand for 0.000000000000000001!*

As we were able to probe successively smaller distance scales we found a sequence of substructures within substructures, however since the 1960s when probes of the proton revealed their quark substructure there has not been any further evidence for another layer of substructure below that of quarks and electrons. We don't know that the things we call elementary particles (like quarks and electrons) even have a size, we only know that at the distance scales we've probed we don't see any structure, so if they do have a size it's smaller than 10^{-18} meters. In other words, we haven't been able to resolve any structure inside them down to the distance scale probed by the Large Hadron

Collider, which is currently the scientific instrument that can probe the shortest distance scales. We'll discuss the Large Hadron Collider more in the Sect. 7.3.

7.2 Atoms

When scientists show pictures of atoms they often look like miniature solar systems with the nucleus playing the role of the sun and electrons playing the role of planets orbiting around the nucleus in concentric circles, see Fig. 7.2.

▷▷▷ **Misconception !!** Atoms are tiny Solar Systems.

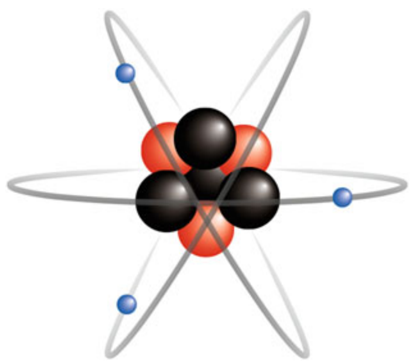


Fig. 7.2 The common picture of an atom shows electrons orbiting a nucleus. In this case, with three electrons and three protons, the picture represents a Lithium atom.

But the electrons inside atoms don't really follow orbits, they don't really move along well-defined paths (see Sect. 2.3). If we repeatedly tried to measure the position of an an electron in a hydrogen atom that is prepared identically thousands of times, we would find a distribution of positions, just like we would find a distribution of locations of people inside a country. The strange thing is that the electron distribution is for just a single electron, so we can't associate a well-defined position or even an orbit for the

The probability is calculated from the square of the wavefunction, see Chap. refduality.

electron. Quantum mechanics only allows us to calculate the probability of the electron being at some position inside the atom. This distribution of positions inside the atom depends on the energy that the electron has. If we give the electron slightly more energy it can have a completely different distribution of positions in space, see Figs. 7.3, 7.4, and 7.5.

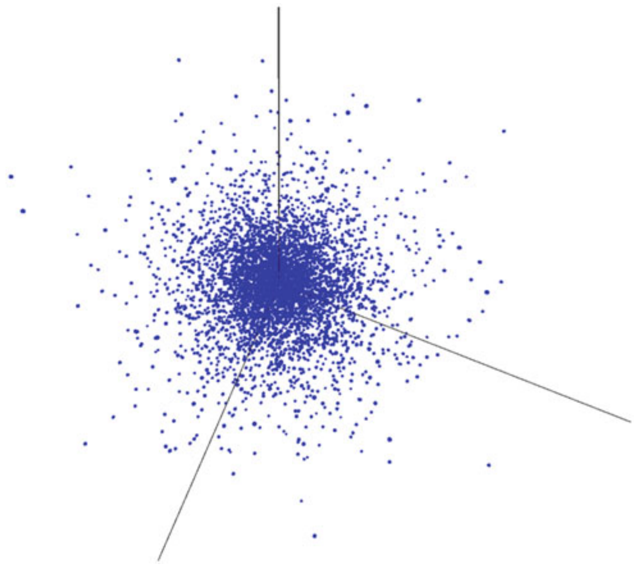


Fig. 7.3 A computer simulation of the measured position of the electron in 5000 identically prepared hydrogen atoms (in the lowest energy state). *Each dot* corresponds to one measured position of one electron in one atom. The lines represent the three spatial directions.

The nucleus of a hydrogen atom is just a single proton.

In the lowest energy state the electron in a hydrogen atom has a spherically symmetric distribution close to the nucleus. In the second lowest energy state, Fig. 7.4, the electron is further from the nucleus but curiously the probability distribution is far from spherically symmetric, it has a sort-of dumbbell shape.

▷▷▷ **More Accurate !!** The distribution of the electron also depends on the angular momentum in addition to the energy.

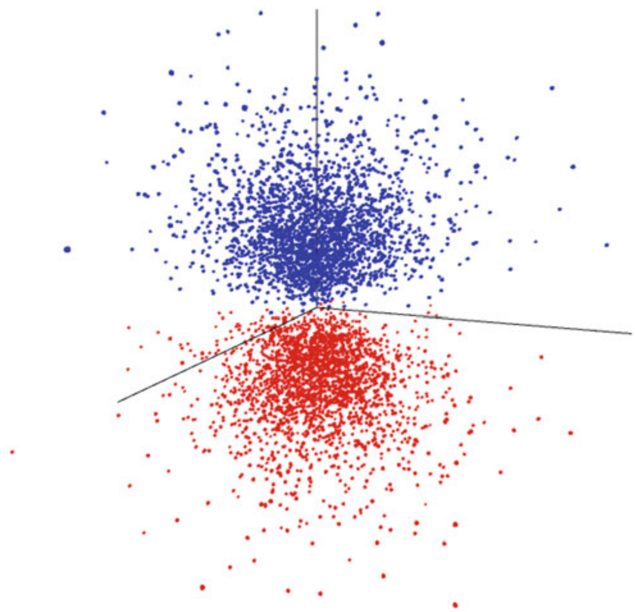


Fig. 7.4 The (computer simulated) measured position of the electron in 5000 identically prepared hydrogen atoms (in the second lowest energy state). *Each dot* corresponds to one measured position of one electron in one atom. This is not to scale with the previous figure. The lines represent the three spatial directions.

In the third lowest energy state, Fig. 7.5, the electron is even further from the nucleus, and the probability distribution has an even more elaborate shape with an extra ring around the middle of a dumbbell. As we continue to go to higher energy levels, or more complex atoms with many electrons, we find increasing complicated electron distributions. The simple picture of a miniature solar system is certainly easier to remember (and to draw!) but is doesn't really do justice to the real beauty of atoms.

The coloring in these electron "cloud" pictures is related to the complex phase of the wavefunction.

You may have wondered why I kept referring to "5000 identically prepared hydrogen atoms" over and over again. The reason is that if we actually tried to do the measurement of the electron position we would give so much energy to the electron it would knock it completely out of the atom. This is the uncertainty principle at work! We will explore this in more detail in the next section.

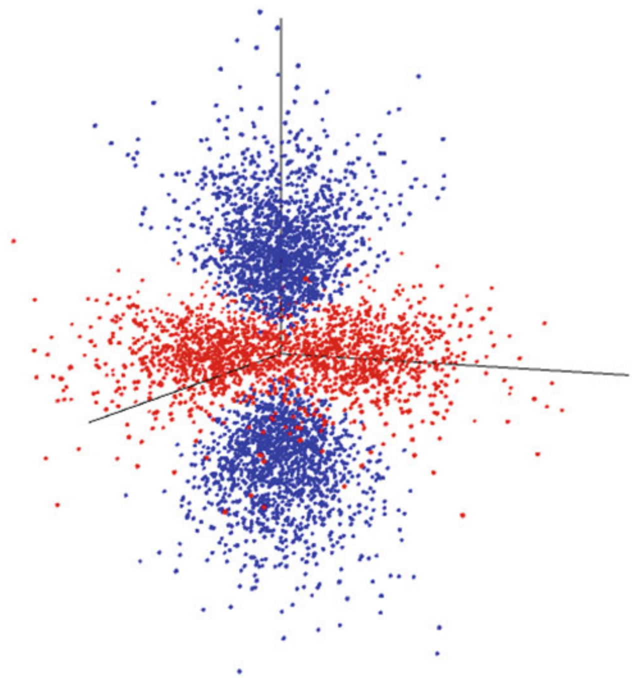


Fig. 7.5 The (simulated) measured position of the electron in 5000 identically prepared hydrogen atoms (in the third lowest energy state). *Each dot* corresponds to one measured position of one electron in one atom. This is not to scale with the previous two figures. The lines represent the three spatial directions.

7.3 Seeing Atoms

How do we see the positions of electrons inside atoms? Many people would suspect that since atoms are so small we would just need to build a very large optical microscope. The reasoning being that that the bigger the microscope the more magnification power it has (Fig. 7.6).

▷▷▷ **Misconception !!** Seeing atoms requires a really big optical microscope.

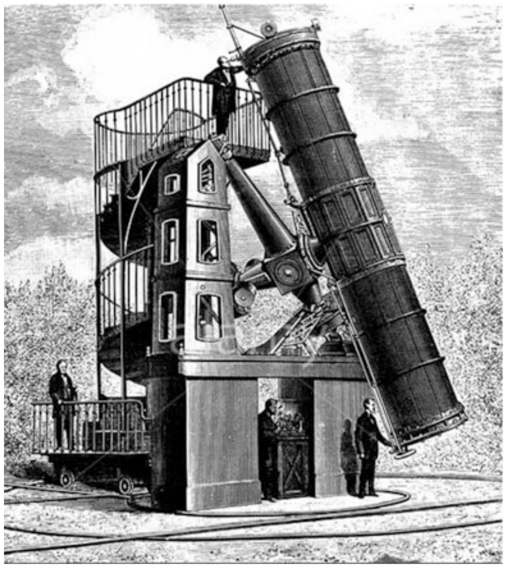


Fig. 7.6 A really big microscope? This is actually the telescope of the Paris Observatory, circa 1870.

The problem with this idea is that light can behave like both a wave and a particle (see Chap. 2). That light can behave like a wave has been known since the time of Thomas Young, who first performed the famous double slit experiment. To do his experiment Thomas Young needed a light source that acted like a single point source. That's because to see the effect he was looking for he needed to start with a lightwave where all the crests of the waves were orderly and lined up. You can imagine the difference between looking at the waves from a single stone dropped in a pond and the waves from many stones dropped all over the pond. Far away from where the single stone was dropped the wave crests are almost perfectly parallel and uniformly spaced. In the pond where many stones were dropped all over, no matter where we are the wave crests are a higgledy-piggledy mess. We call a light source that produces nice uniform waves a "coherent" light source. To make a coherent light source Thomas Young let sunlight go through very small hole, this was like the pond with a single stone. He let the light from the single hole pass through to two very narrow, closely spaced slits. Since the coherent light source he used was just a small hole, the light was very faint and the experiment had to be conducted in a dark room. Nowadays

Thomas Young (1773-1829) the English polymath who first showed that light can behave like and wave and helped to decipher the Egyptian hieroglyphs on the Rosetta Stone.

it's very easy to reproduce his experiment since we can use lasers which are very powerful coherent light sources. Passing laser light through two closely spaced slits and letting it fall on the screen we see a pattern of dark and bright bands. This is how Young knew that light could behave like a wave. The bright bands correspond to places where the successive crests of waves meet up and the dark bands correspond to places where the crest meets the trough and the waves cancel (Figs. 7.7 and 7.8).

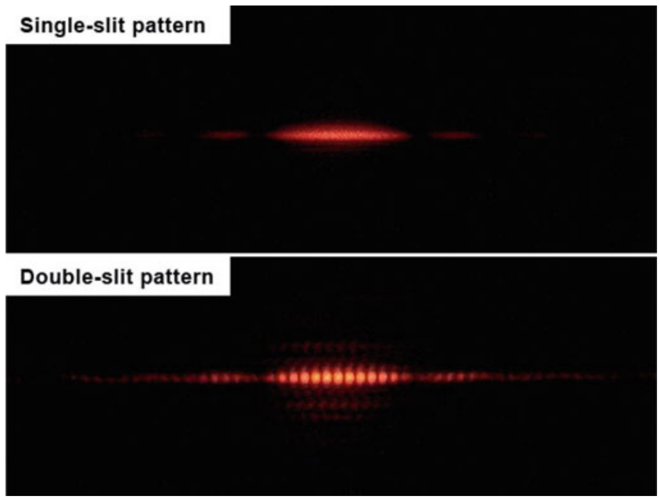


Fig. 7.7 Laser light going through a single slit and through two identical slits 0.7 millimeters apart. The alternating *dark* and *bright* fringes in the *bottom photo* are clear evidence of wave behavior; compare with Fig. 3.5.

Fig. 7.8 The wavelength of a wave is the distance from point on the wave to the next identical point.

We call the distance between the crest and the successive crest (or from trough to trough) the wavelength. You probably know that a sound wave with a certain wavelength has a corresponding frequency. The short wavelength sounds that come from a small tweeter are the high-frequency sounds, the long wavelength sounds from a large base speaker are the low-frequency sounds. This is true for any type of wave—be it light waves, sound waves, seismic waves, or water waves—the highest frequencies correspond to the shortest wavelength and lowest frequencies correspond the longest wavelengths.

A very low pitched sound at 20 oscillations per second has a wavelength around 17 meters, while a very high pitched sound at 20,000 oscillations per second has a wavelength around 0.017 meters (about 2/3 of an inch).

However we also know that light can behave like a particle. One of Einstein's many contributions to science was his unraveling of the photoelectric effect. When we shine light onto a metal, electrons can be knocked out, but the energy that the electrons have when they emerge from the metal doesn't depend on the intensity of the light it depends on the frequency of the light. The highest frequency light produces the highest energy electrons. Einstein realized that this meant that the energy in the light must be carried by individual energy packets, photons, and that each photon's energy was proportional to its frequency. Einstein's explanation was one of the earliest examples of how confusing quantum mechanics can be. In the double slit experiment we see that light can act like a wave and in the photoelectric effect where the light interacts with electrons directly we see that light can act like a particle. Maybe even more confusing is that the energy of the particle is determined by frequency of the wave. This makes no sense on the everyday scale of humans, but that is simply because we don't experience things at the subatomic scale.

Einstein actually won his Nobel prize for explaining the photoelectric effect, relativity was considered to radical at that time.

Now we can get back to our problem with the really large microscope. In order to see something very, very small we need to use light with a wavelength that is smaller than the size of the object we are trying to study. That means that the photons in the light beam will have a large frequency, and hence a large energy. To get a sense of the energies involved let's look at some examples. The wavelength of the radio wave used for Wi-Fi is about 5 inches (about 13 centimeters). The only difference between radio waves and light is that radio waves have much longer wavelengths than visible light. The energy of an individual photon used in Wi-Fi is about one 100,000th (or 10^{-5}) of an electron-volt. One electron-volt is the typical energy involved in atomic transitions, that is why the energy that a

Trying to see a small particle with a particular wavelength of light is analogous to waves bouncing off of boats. A wave will reflect back off of a cruise ship if its wavelength is short compared to the size of the ship, but if the wavelength is much longer than the ship, the ship will just ride over the wave and there will be no reflection.

standard AA battery gives to a single electron is about 1.5 electron-volts, so we call it a 1.5 volt battery. Visible light has much shorter wavelengths and higher photon energies. If we go up to ultraviolet light, which has a wavelength slightly shorter than visible light, the photons have an energy of about five electron-volts corresponding to a wavelength of about one 100,000th of an inch (about 3×10^{-7} meters). But this is still too long a wavelength to resolve an atom, since atoms are about 4×10^{-9} inches (or 10^{-10} meters). So it is impossible to see an atom with visible light (Fig. 7.9).

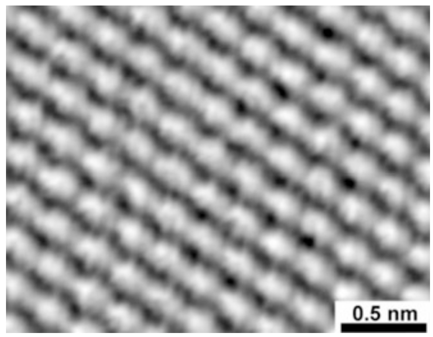


Fig. 7.9 Image of carbon atoms in graphite (pencil lead) obtained with a scanning tunneling electron microscope. The *inset* scale is 0.5 nanometers (about 2×10^{-8} inches).

In order to make a microscope that can resolve individual atoms we need a wavelength smaller than one 100 millionth of an inch, which corresponds to an energy of about 5000 electron-volts. It is somewhat difficult to get such high-energy photons (these are X-rays) and even more difficult to build a lens capable of focusing X-rays in order to make an image. However we can produce and focus high-energy electrons, so we can make an electron microscope. Electrons can also act both as a wave and a particle so it is perfectly reasonable to use them instead of photons, but we're not seeing things directly with such a microscope. We would not be seeing things directly with our eyes even if we used X-rays; our eyes cannot see X-rays. We can record where the electrons are scattered after hitting our target atoms and then infer what they scattered off, so this is seeing in a generalized sense.

Ernest Rutherford (1871-1937)

If we keep going higher in energy we can keep probing smaller scales. Ernest Rutherford (along with Hans Geiger and Ernest Marsden) discovered the nucleus of the atom by sending a beam of alpha particles at a thin layer of gold foil. Rutherford wanted to see how much the positively charged alpha particles could be deflected by the atoms. Since the alpha particles are about 8000 times heavier than an electron, an interaction with an electron could not deflect the alpha particle: it would be like a bowling ball colliding with a ping-pong ball, the bowling ball will continue on unaffected. However electrons make up less than one 2000th of an atom's mass, and Rutherford was interested in the effects of the mysterious remainder that made up the bulk of an atoms mass. The prevailing picture at the time was that this bulk was made up of a soothly distributed positive charge.

Geiger went on to fame as the inventor of the Geiger counter.

The alpha particle is composed of two protons and two neutrons tightly bound together, which is just the nucleus of a helium atom. Since it is so tightly bound it is often produced in nuclear decays, when unstable nuclei break apart into smaller pieces.

In order to detect the scattered alpha particles they set up a small screen coated with zinc sulfide. When an alpha particle hit the screen it would produce a small flash that could be detected by eye when looking through a low-power microscope in a darkened room. This was tedious and frustrating work. Marsden and Rutherford took turns, one looking through the microscope while the other wrote down the results. They found that, unexpectedly, the alpha particles could be deflected by a large angle, sometimes almost straight back the way they came. The alpha particles had energies around 5,000,000 electron-volts, allowing them to probe down to 10^{-14} meters. Rutherford calculated that to bounce back like that the alpha particle had to come within 10^{-14} meters of the total positive charge. This is a distance 10,000 times smaller than the size of the atom. This was the evidence that led Rutherford to realize that positive charges inside the atom we tightly concentrated in the nucleus rather than spread out as had been previously thought.

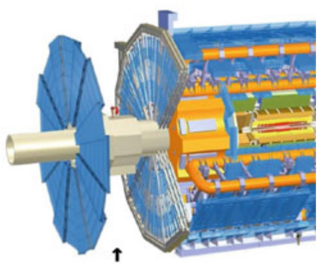
A similar story was repeated 50 years later, when an experiment at the Stanford Linear Accelerator Center in association with physicists from MIT discovered quarks. The accelerator produced a beam of electrons with 20 billion (as in 20,000,000,000) electron-volts each that was send into liquid hydrogen. Sophisticated (for the time) automatic electron detectors could be moved to different positions around the liquid hydrogen target, and the team found that sometimes the electrons lost a lot of energy and were scattered at large angles. This meant that the electric charge inside a proton was not spread out smoothly but concentrated in small regions. Most theoretical physicists were initially sur-

Jerome Friedman, Henry Kendall, and Richard Taylor won the 1990 Nobel Prize in Physics for leading this experiment that provided direct evidence for quarks.

prised, but lead by James D. Bjorken and Richard Feynman they were gradually convinced of the existence of quarks. Further confirmation came from CERN, the European Organization for Nuclear Research, where similar experiments were done using neutrinos instead of electrons. It was eventually established that protons and neutrons have three quarks inside each of them.

Neutrinos are very light and have no electric charge, so they only interact with quarks via weak interactions.

The cutting edge of this kind of experiment is the Large Hadron Collider at CERN, near Geneva Switzerland, which uses generalizations of these basic techniques. A technical complication is that rather then sending a beam towards a target, at the Large Hadron Collider two beams of protons have to collide almost head-on. This requires a very complicated sequence of magnets to focus each beam of particles down to as small a size as possible. The advantage is that this set-up maximizes the amount of energy available in the collision. The experimental teams of thousands of people try to understand what happens in these collisions by recording the directions and energies of the different particles that come out of each collision. Rather than having a movable detector, the collision region is completely surrounded by several layers of different kinds of detectors, that are optimized for finding different types of particles. The detectors are as large as a four story building (Fig. 7.10).



Drawing showing half of the ATLAS detector, one of the two main detectors at the Large Hadron Collider. The arrow indicates the height of an average human.

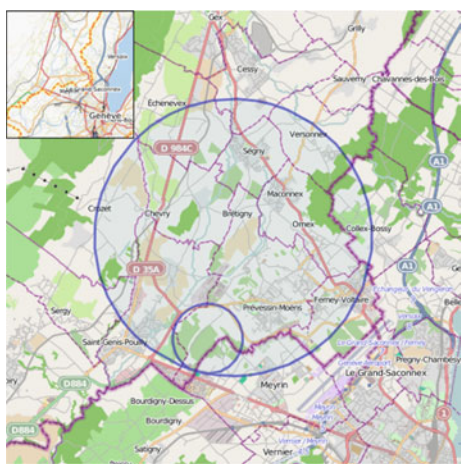


Fig. 7.10 Map of the Large Hadron Collider (large ring) which goes under several towns and villages. I lived in one of them, Ferney-Voltaire, for seven months while visiting the lab.

The main difference between the classic experiments described above and the Large Hadron Collider is that we now use particles with much higher energies: the protons each have energies around 6.5 trillion (6,500,000,000,000) electron-volts. To reach such high energies the ring that accelerates the protons and stores the beams is about 9 kilometers (5 miles) across!

▷▷▷ **More Accurate !!** The Large Hadron Collider is our most advanced instrument for probing structure at small scales. It is allowing us to study particle interactions at a scale of 10^{-18} meters, which is 100,000,000 times smaller than an atom!

Chapter 8

Mini Black Holes

8.1 Black Holes

▷▷▷ **Misconception !!** The Large Hadron Collider could spawn a planet devouring mini black hole.

Around the time that the Large Hadron Collider first turned on in 2008, they were a series of dramatic headlines claiming that the Large Hadron Collider could produce mini black holes and that these black holes would then eat up the Earth! These stories were all over the internet, including sites that you would've thought were more reliable, serious sources. For example, the *National Geographic News* on September 10, 2008, had the following headline: "Worst Case: Collider Spawns Planet-Devouring Black Hole." (Fig. 8.1).

Of course the people who built the Large Hadron Collider had no intention of blowing themselves up or destroying the Earth that they live on. They knew that there would not be a problem with running the Large Hadron Collider because the amount of energy per collision that we can currently produce in an experiment is minuscule compared to

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_8

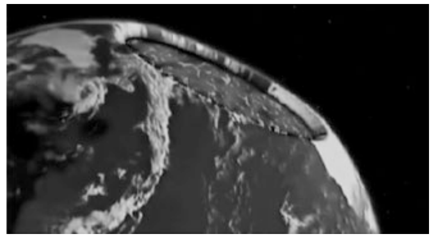


Fig. 8.1 Artist's impression of a mini black hole devouring the Earth, starting out in Geneva, Switzerland.

the energies that are produced in outer space. For example, cosmic ray showers hit the Earth constantly, and some of those collisions have energies much larger than those that occur at the Large Hadron Collider. Collisions from cosmic rays with comparable energies to the Large Hadron Collider happen once per square meter per year on the surface of the Earth. Cosmic rays with energies a thousand times larger arrive at a rate of one per square kilometer per year. Nature has been running this experiment for 4 billion years with no catastrophic production of black holes. In addition cosmic rays have been hitting the moon for a similar amount of time and, more importantly, lots of other very dense objects like white dwarfs and neutron stars. Even if the mini black holes could be produced in cosmic rays and then punch right through the Earth, they could not punch through such dense objects. The fact that we have seen lots of white dwarfs and neutron stars meant that the Large Hadron Collider was safe to turn on.

On the Web: `Large Hadron Collider Safety Assessment: bit.ly/2ckKnT3`.

A more interesting question is why would we expect that mini black holes could be produced in the first place. In the standard theory of gravity, Einstein's theory of general relativity, they would not be produced at the Large Hadron Collider. Einstein's theory does predict black holes: whenever a sufficiently large mass (or energy) is compressed in a small enough region, a black hole should form. For example, if the mass of the Earth (about 6×10^{24} kilograms) was compressed in a sphere of radius 8.8 millimeters (about one third of an inch), it would form a black hole. The basic idea of a black hole is just that with a sufficiently strong gravitational field even light cannot escape gravity's pull. This idea goes way back before Einstein, starting with John Michell in 1784. You may have heard that the escape veloc-

On the Web: `Michell's paper: bit.ly/2cXG9yu`.

ity for a rocket to leave the Earth is about 11.2 kilometers per second (about 7 miles per second or 33 times the speed of sound), but with all of its mass compressed into a radius of 8.8 millimeters, the escape velocity of the Earth is greater than the speed of light, and so we have a black hole.

The speed of light is 3×10^8 *meters per second or 186,000 miles per second.*

The radius that a mass must be squished down to in order to form a black hole is called the object's Schwarzschild radius, after Karl Schwarzschild, who was the first person to find a solution of Einstein's equations describing a spherical black hole. Anything venturing inside of a black hole's Schwarzschild radius will not get back out. We believe that one way to achieve the super-high densities needed to form a black hole is during certain very large supernova explosions. It is also possible that some black holes formed when the universe was very young, before there were any galaxies or stars; these are called primordial black holes. Despite our fuzzy knowledge about how black holes are made, we know that they are not just theoretical daydreams: we have plenty of evidence that black holes really exist. There is a 4.1 million solar mass (or 8.2×10^{36} kilograms) black hole at the center of our Milky Way galaxy. Astronomers have watched for years as stars zip in tight orbits around a region that looks quite empty, except for some occasional flares (possible these occur when something falls into the black hole). This black hole should have a Schwarzschild radius of 1.2×10^{10} meters (about four times the radius of Uranus's orbit).

▷▷▷ **More Accurate!!** More generally the point of no return when approaching a black hole is called the event horizon; for a black hole that isn't rotating, the event horizon is at the Schwarzschild radius. The event horizon of a rotating black hole is more complicated than a sphere.

More recently the LIGO experiment has recorded the gravitational waves produced by the collision and merger of two black holes with masses about 30 times that of the Sun. Each of them would have a Schwarzschild radius of about 90 kilometers (55 miles). The astounding feat of capturing the tiny ripples in space from a collision over one billion light years away (a light year is the distance light travels in one year, 9.5×10^{12} kilometers or 5.9×10^{12} miles) surprised a lot of people. Since it took the gravitational waves traveling at the speed of light over a billion years to get here, it also

On the Web: `Gravitational waves from black hole merger: bit.ly/2deUF7e`.

Max Planck (1858–1947) introduced a constant that appears in all quantum mechanical formulas, it is now known as Planck's constant.

On the Web: `Short distance gravity tests: bit.ly/2cs2HYK`.

gave the scientists an opportunity to joke that the collision they observed "happened a long time ago in a galaxy far, far away!"

You have probably noticed a trend in the numbers I've given for black holes. The larger the mass of the black hole, the larger the Schwarzschild radius. In fact, the relationship is strictly linear; if we double the mass, we also double the Schwarzschild radius. With the energy available in our colliders, this means we are talking about very small back holes. The typical energy in a collision at the Large Hadron Collider is 1 tera electron-volt (10^{12} electron-volts). If this energy was sufficiently compressed to form a black hole, then Einstein's theory would predict it to have a Schwarzschild radius of 3×10^{-51} meters. Talking about such tiny black holes, we might worry that we have extrapolated Einstein's theory beyond where is applicable. Einstein's theory of gravity is notoriously incompatible with quantum mechanics, and a simple estimate shows that quantum effects should become important for gravity at distance scales around the Planck length: 10^{-35} meters. That is a distance that is 10^{20} times smaller than the size of a proton! We are nowhere near testing gravity at such length scales. Currently precision tests of gravity go down to around 10^{-4} meters.

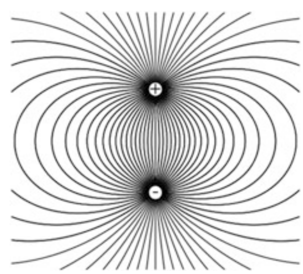
8.2 Black Holes and Extra Dimensions

Since physics is really an experimental science, the fact that we haven't tested gravity at very small scales means that Einstein's theory of gravity might break down before we get to the Planck length. In principle it could break down at any length scale smaller than 10^{-4} meters. This is exactly what happens in theories of "large" extra dimensions, where large here means large compared to 10^{-19} meters, the length scale that we can probe with the Large Hadron Collider. To see why, we need to know a little more about how gravity works.

Newton famously showed that the force of gravity between two masses is proportional to the product of the masses and falls off as one over the square of the distance between them. That is, if we double the distance, the force of gravity is reduced by a factor of four. There is a constant multiplying all of this that gives the correct units of force, it is called Newton's constant. In our modern language, Newton's constant is proportional to the square of the Planck length. One way of looking at it is that gravity is so weak compared to the other forces because the Planck length is so small.

The fact that gravitational forces (like electric forces between two charges) fall with one over the distance squared is a consequence of living in a three-dimensional space. If we lived in a four-dimensional space, the force would fall like one over the cube of the distance. The distance in the denominator is raised to a power, and the power is always the number of spatial dimensions minus one. The reason for this is that the gravitational field has to spread out as we go farther from a massive object: in a two-dimensional world, it would spread over a circle whose circumference grows linearly with distance; in a three-dimensional world, it spreads out over a sphere whose area grows with the square of the distance. An easy way to picture this is using "field lines," a concept introduced by Michael Faraday to help understand electromagnetic forces. Faraday suggested that it is helpful to think of the forces between electric charges as the result of fields that pervade space. An electric charge produces a field around it, and other electric charges respond to that field. James Clerk Maxwell later used these ideas to show that light waves are oscillations of these fields. To visualize this, we can draw simple diagrams of the fields. For each electric charge, we draw evenly distributed lines coming out it in all directions. Larger charges have more lines; double the electric charge means double the number of electric field lines. The direction of the field lines indicates the direction that a very small charge would move at that position. Like charges repel, so the small charge would move outward along the field line, while an opposite charge would move inward toward where the field lines are concentrated. The density of the field lines indicates the strength of the electric field. The same idea works for gravity where the number of field lines is proportional to mass (or energy) rather than charge, but there is no gravitational repulsion, only attraction. Figure 8.2 shows how gravitational field lines would work in two and three dimensions.

Michael Faraday FRS (1791—1867) used his intuitive understanding of field lines to discover electromagnetic induction which lead to electric motors and generators.



The field lines between a positive and a negative charge.

Now suppose there were some small extra dimensions (imagine an extra circle or sphere at each point in ordinary space) that we can't move through either because they are wrapped too tightly or we are simply stuck at a special point. If we are always stuck at one point in the extra dimensions, then we won't be able notice them easily, and/or if they are much smaller than 10^{-4} meters, then they would

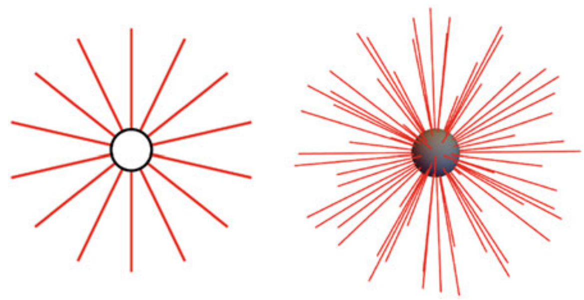
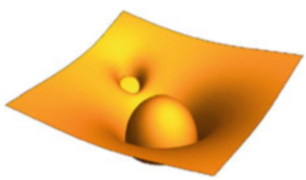


Fig. 8.2 Field lines in a two-dimensional world and a three-dimensional world. The field lines spread faster in higher dimensions.



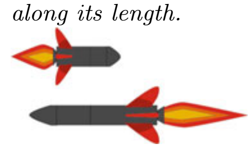
In Einstein's theory of gravity the orbits of planets can be understood as the planets following the shortest path in the curved spacetime around a star.

also be extremely hard to discover. But even if we can't move in the extra dimensions, gravity will, since, as Einstein taught us, gravity is the curvature of space and time, so gravity can go wherever there is space and time, including in the extra dimensions. In that case the way the gravitational field lines spread out will change from our standard picture of gravity. When we look at distances smaller than the size of the extra dimensions, then the gravitational field falls off with a larger power of distance because there are more dimensions to expand into. But when we look at distances larger than the size of the extra dimension, everything will work as we would expect for three spatial dimensions, and the gravitational field will fall like one over the square of the distance, just as Newton thought.

Turning it around and starting from large distances and going to smaller distances, we would see that at first the gravitational field grows as usual, but when we get to distances smaller that the size of the extra dimensions, it starts growing much faster. Depending on the sizes and numbers of extra dimensions, it is quite possible that gravity becomes strong not at the Planck length but at distances like 10^{-18} meters. In other words the hidden extra dimensions would be masking the fact that the true Planck length is really 10^{-18} meters and not the 10^{-35} meters that we naively thought. This may still seem like an incredibly tiny distance, but this is the scale that the Large Hadron Collider can probe. If we lived in such a universe, then it would be possible to create mini black holes at the Large Hadron Collider!

It still may seem a stretch to you that colliding two ordinary protons could produce the extreme densities needed to produce a black hole. It is perhaps easier to think about how things work if we imaging moving at almost the same speed as one of the protons. During the collision we would see this very slowly moving proton being approached by another proton that is moving almost at the speed of light. Since it is moving near the speed of light, we know that we need to take into account Einstein's special theory of relativity. Things moving near the speed of light do not behave like the things we see in everyday life. For this situation the most important thing to focus on is the apparent length contraction. The proton will appear to be very squished along its direction of travel. This does not just apply to the proton but also its electric and gravitational fields. Figure 8.3 shows what happens to the electric field lines for protons at three different velocities. As the proton moves faster, the field lines are squished more and more so that they are denser in the directions at right angles to the direction of motion. The density of the field lines tells about the strength of the field, and as we see in Fig. 8.3, in the regions where there are few field lines, the electron travels in almost a straight line, and in the region with a large number of field lines, the electron's path is curved.

From the point of view of the top rocket, the identical rocket speeding by below in the opposite direction looks shrunk along its length.



From the point of view of the bottom rocket, it is the top rocket that is shrunk. Special relativity makes both points of view consistent because events that seem simultaneous from one view point are not simultaneous in the other.

The gravitational field lines are squished just like the electromagnetic field lines, but there is an extra twist, the strength of the gravitational field depends on the total energy of the proton, and the faster it travels, the more energy it has, and the stronger the gravitational field. At high enough energies the field is so strong that even if initially the protons looked as if they would pass by each other, the strong gravitational scattering can make them pass very close together, so close that the fast proton is within the Schwarzschild radius of the slow proton. This happens with a large probability only when the energy in the collision is high enough to resolve the Planck length. In an ordinary world with three spatial dimensions and a Planck length of 10^{-35} meters, this would take a tremendous amount of

On the Web: `Black hole creation in colliders: bit.ly/2cJPDig`.

energy, about 2×10^{18} times as much energy as the proton mass. In a world with extra dimensions and a Planck length of 10^{-18} meters, this can happen with an energy around a thousand times as much as the proton mass, which is the energy available in a collision at the Large Hadron Collider. So if we see mini black holes being produced at the Large Hadron Collider, then we will know that gravity is very different at the scale of 10^{-18} meters than it is in everyday life.

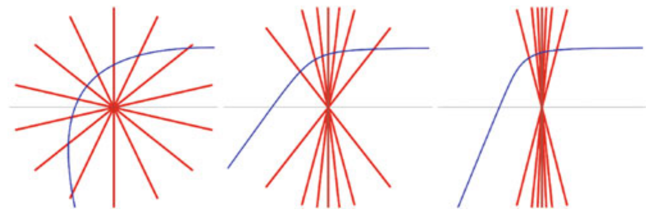


Fig. 8.3 (*Left*) An electron (*curved line*) scattering through the electromagnetic field of a proton. (*Center*) An electron scattering through the electromagnetic field of a proton moving to the right at 94% of the speed of light (the motion of the proton is subtracted out, so it is at a fixed position in the diagram and only the electron is shown moving). (*Right*) An electron scattering through the electromagnetic field of a proton moving at 99% of the speed of light.

Previously I said that we knew the Large Hadron Collider was safe to turn on because cosmic rays had not produced disastrous effects on Earth and on neutron stars, doesn't this mean that we already know that mini black holes will not be produced? Note quite, it means that if mini black holes are produced, they will not be disastrous! This seems very counter-intuitive since black holes have a reputation for swallowing up everything they get near; however it has been known since the 1970s that quantum mechanics makes lightweight black holes evaporate. By a "lightweight" black hole, I mean one that has a Schwarzschild radius that is not too much larger than the Planck length. The calculation that shows how black holes can evaporate combines quantum mechanics, and gravity is a subtle way and was first done by Stephen Hawking. It was this calculation that made him a star in the physics world. The essence of Hawking's calculation is that, because of quantum mechanics, pairs of particles (e.g., an electron and and antielectron) can just appear out of empty space. This is because there is an uncertainty principle (see Sect. 2.3) relating energy and time, which means that energy con-

Antielectrons are also called positrons. They were predicted by Paul Dirac in 1928 as a consequence of combining special relativity and quantum mechanics and discovered in 1932 by Carl Anderson, who won the Nobel Prize for his discovery in 1936. Anderson also discovered the muon.

servation can be violated briefly if it only happens for a short enough time. The bigger the violation, the shorter the amount of time that it is allowed. This means that particles are constantly appearing out of nowhere and disappearing before we have time to notice them directly. However if a particle and antiparticle pair materialize near a black hole, it is possible for one of them to fall into the black hole while the other one escapes. The particle that falls in can never escape to reunite with its partner, which is then free to speed off far away from the black hole. Energy is conserved in the end because the mass of the black hole decreases by the amount of equivalent energy that the escaping particle took away. Thus the black hole evaporates by emitting particles. For a very heavy black hole, this process is very slow and can easily be overcome by other bits of matter (or radiation) falling into the black hole. For lighter black holes the evaporation process is faster, and for very light black holes, the "evaporation" is more like an explosion.

We can see the effects of pairs of charged particles appearing and disappearing indirectly through its affects on the strength of electric attraction and repulsion at very short distances.

Einstein's formula, $E = m\,c^2$, gives the equivalent energy, E, for a particle at rest with a mass m, where the conversion factor, c^2 is the speed of light squared.

For a mini black hole produced at the Large Hadron Collider, one finds that it would evaporate in as short time as 10^{-22} seconds, so it would not have time to swallow anything. If it was produced, it would almost immediately disappear in a burst of particles (Fig. 8.4).

▷▷▷ **More Accurate !!** Black holes are not completely black; they emit Hawking radiation. The smaller the black hole, the faster it emits Hawking radiation.

Fig. 8.4 Stephen Hawking enjoying zero gravity aboard the "vomit comet," a modified Boeing 727 in free fall.

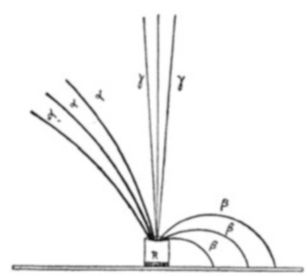
Chapter 9

Particle Colliders and the Universe

9.1 The Higgs Boson

▷▷▷ **Misconception !!** The Higgs Boson could wipe out the Universe.

Another scary headline appeared on *CNET* in 2014 proclaiming "Stephen Hawking: God particle could wipe out the Universe." To understand what Hawking was referring to takes a little bit of work. First of all, what the media calls "the God particle" is what physicists call the Higgs boson. The Higgs boson is a special type of particle that was discovered (to great fanfare) in 2012 at the Large Hadron Collider. In some sense this discovery was the most important since the time of Ernest Rutherford. Among other things, Rutherford was one of the first scientists to study radioactivity. In one of his early papers, he classified the three different types of radiation he had found. He called them alpha rays, beta rays, and gamma rays (also written as α, β, γ, the "ABCs" of the Greek alphabet).



Rutherford's drawing showing how the radiation coming from a radioactive sample could be separated by a magnetic field. Here alphas are deflected to the left, βs are deflected much more and in the opposite direction, while the γs are unaffected.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*, Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_9

Rutherford later showed that the rays he had called alphas were in fact made up of the nuclei of helium atoms. These alphas can be spit out of a nucleus in some types of radioactive decays. With our more modern nomenclature we refer to the class of particles made of quarks as hadrons. The Large Hadron Collider has hadron in its name because it collides protons together, and protons are a type of hadron, being made out of three quarks. The alpha is made up of two protons and two neutrons, so it is a particular arrangement of 12 quarks.

There are six types of quarks that have somewhat whimsical names. Up, charm, and top have a charge equal to two thirds of a proton charge, while down, strange, and bottom have negative charges, one third of the charge of an electron. A proton is made of two up quarks and one down quark, while a neutron is made of one up quark and two down quarks.

The particles that Rutherford called betas turned out to simply be electrons. We now know that there are other particles similar to electrons like muons and neutrinos. None of these particles feel the strong force that binds quarks tightly together. We call this category of particles leptons. The names hadron and lepton come from Greek words meaning heavy and light; often particles made out of quarks are much heavier than the leptons.

Finally the rays that Rutherford called gammas were really just very high-energy photons, that is, particles of light. Since Rutherford's time, we've discovered other particles that are similar to photons; these are the three types of "force carriers." Electrically charged particles like electrons can feel the electromagnetic force because they can exchange photons with other electrically charged particles. In addition to the photon, there is also the gluon which carries the strong force that binds quarks together to form hadrons (the name gluon was actually inspired by the requirement for some very strong "glue" to hold the quarks together). There are also W and Z bosons associated with the weak force that is responsible for the type of radioactive decay where electrons (or antielectrons) and neutrinos are emitted by the nucleus, and the fusion process in the Sun where protons are converted to neutrons as heavier elements are built up starting with only hydrogen. Physicists call all of these force-carrying particles gauge bosons. Amazingly Rutherford knew about the three classes of particles that make up almost all of the modern "standard model" of particle physics: hadrons, leptons, and gauge bosons.

The standard model specifies which particles can interact with which, and the strength of each interaction. Applying the combination of quantum mechanics and special relativity (known as quantum field theory) allows us to make very precise predictions that have been verified in an amazing range of experiments.

While the gluon and photon are massless, the W has a mass about 80 times as large as a proton mass, while the Z has a mass about 90 times as large as a proton mass. This difference in masses is tied to how they interact differently with the Higgs boson.

What was later learned through careful experiments is that there is a fundamental difference between the categories: quarks and leptons have spin one half, while gauge bosons have spin one. The word spin sounds like it refers to some type of rotational motion, but it is quantum mechanical so a particle with spin does not quite behave like a spinning top. As far as we know quarks and leptons don't occupy any volume, so there is "nothing to spin."

For an example of how a quantum spin is different from a rotating ball, consider a spinning top with an electric charge. Such a charged top would be deflected if it went through a varying magnetic field. The amount of deflection would depend on the angle between the direction of the magnetic field and the direction of the axis that the top is rotating around. If we sent a "beam" with many such electrically charged tops through a magnetic field, and each top rotated around an axis in a random direction, the beam would spread out since each top would get a different deflection. However if we send an electron through a varying magnetic field, we get only two types of deflections, corresponding to the electron spin direction aligned with the magnetic field and the electron spin anti-aligned to the magnetic field. This is how Otto Stern and Walther Gerlach first demonstrated that electrons come in two possible spin states rather than a continuous distribution of different spin directions. This is a general property of particles with spin one half. Massive spin-one particles like the W and Z come in three spin states. In general the number of spin states of a massive particle is two times the spin plus 1.

Particles with a spin that is a whole number are generally called bosons, while particles with spin that is half of an integer are called fermions. The names were invented by Paul Dirac as homages to Satyendra Bose and Enrico Fermi who performed early studies of the quantum mechanical implications of spin.

There is one type of particle that we've seen in nature that Rutherford had no example of, and that is a particle without any spin, that is, spin zero. The only (possibly) elementary particle that we know of which has spin zero is the Higgs boson. In Rutherford's language we have finally, after more than 100 years, found a new type of particle that Rutherford probably would have called the "delta" (δ, the fourth letter in the Greek alphabet).

Other than its novelty, the reason why scientists are excited about the Higgs boson discovery is that in the standard model, the Higgs boson is required for an explanation of how particles get mass. In rough terms, the mass of a particle is proportional to how much it sticks to "empty" space. However by "empty" space, we don't really mean space that is absolutely empty; we mean it is empty of particles like electrons and quarks, but there is still a Higgs field present. The Higgs boson which was discovered is a fluctuation of the Higgs field just like photons are fluctuations of the electromagnetic field. In "empty," very far from any charged particles, there shouldn't be any electric or magnetic fields either, but the standard model predicts there is a nonzero value of the Higgs field.

To see how this strange state of affairs comes about it is useful to remember that it costs energy to produce an electric or magnetic field. Anybody who pays a utility bill knows that it costs energy and hence money every time they flip a switch to turn on the lights. Light bulbs produce electromagnetic fields; some of the oscillations of the fields are wavelengths that our eyes can detect, which we call light. However in the case of the Higgs field, it costs us energy to turn the Higgs field off.

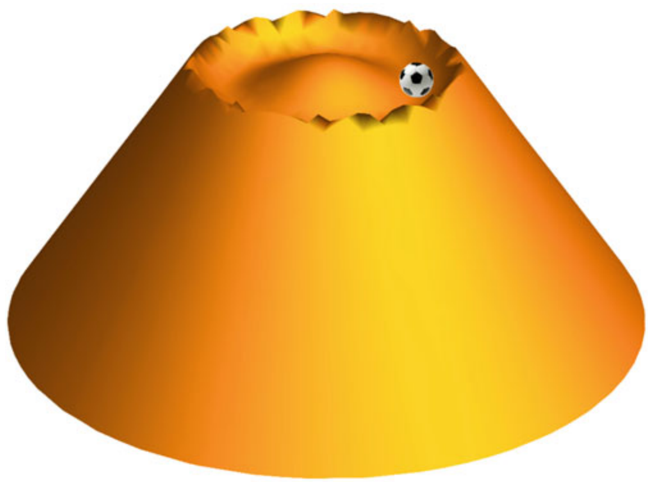


Fig. 9.1 A soccer ball sitting in the caldera of a dormant volcano.

For a simple analogy to the energy requirements for the Higgs field, you can think of rolling a ball around the bottom of the caldera of a volcano as shown in Fig. 9.1. Viewed from above, there is a circular valley around the cone in the center. If we wanted to get the ball out of the volcano, we have to raise its gravitational potential energy to get over the rim. The lowest energy state for the ball is at the bottom of the valley. It doesn't matter where along the valley it sits; if the bottom of the valley is level, there is the same gravitational potential energy at any point. But to get to the middle of the volcano, we have to roll the ball up to the

Gravitational potential energy is a way of keeping track of how much energy is needed to provide the work to lift something in the gravitational field of the Earth.

top of the central peak. We have to do work to push it up that hill; that is how we raise its energy. If we look at a plot of the energy required for a given value of the Higgs field, the central peak of our volcano corresponds to the point where the Higgs field is completely off, and the bottom of the valley corresponds to a nonzero value of the Higgs field, as in Fig. 9.2. If we put our ball down at random, it will roll down to the bottom of the valley. The same is true of the Higgs in the early universe; it may have started out in the off position, but eventually it ended up at the bottom of the valley with a nonzero value.

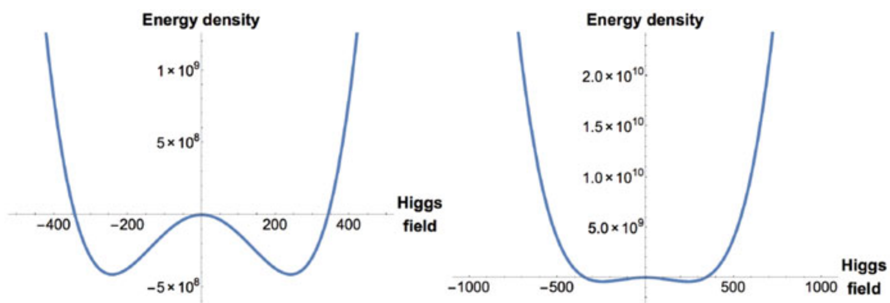


Fig. 9.2 The stored energy density plotted as a function of the value of the Higgs field (in GeV or 10^9 electron-volts). The second plot shows the same thing but over a broader range of values for the Higgs field. In the first plot, we see there is a minimum near 246 GeV, and the energy stored gets larger as we go to slightly larger or smaller values of the Higgs field.

Particles that interact with the Higgs field are slowed down as they move through space. The more that they interact with the Higgs field, the harder it is to push them around and the heavier they are. The photon has no interaction with the Higgs field, since the Higgs boson has no charge, so the photon remains exactly massless. The same is true for the gluon; the Higgs boson does not experience the strong force the way quarks do, so the gluon does not get a mass from the Higgs field. The W and Z bosons do interact with the Higgs, so they do get a mass. The electron also interacts with the Higgs, but this interaction strength is much weaker than that of the W and Z, so the electron is much lighter than the W and Z bosons. The standard model does not explain what the values of the interaction

Of the known particles, the top quark has the strongest interaction with the Higgs, so it has the largest mass, about 175 times the mass of a proton. The Higgs also interacts with itself and gets a mass about 125 times the mass of the proton.

strengths of the quarks and the electron should be. That is one reason why physicists hope to find a more predictive underlying theory. Just as Newton's theory of gravity was a good approximation to Einstein's theory of gravity under certain circumstances, we hope that the standard model is just a useful approximation (that happens to work very well at the length scales that we have probed so far) to a deeper, underlying theory.

Finally we can get to what Stephen Hawking was worrying about! As in the volcano analogy, the energy function for the Higgs field may also have a rim, as shown in Fig. 9.3. In the standard model, we can approximately calculate where the "rim" is. So it's theoretically possible for the Higgs to get into an even lower energy state than at the bottom of the valley, just like the ball going over the rim and rolling down the side of the volcano. If the Higgs field went over the "rim," the masses of all the particles in the universe would suddenly change, and the stable structures that we know about (like atoms and molecules) would no longer exist; they would be turned into some very different structures (or possibly there would be no stable structures at all). So although this transition to a new value of the Higgs field would create a new type of universe, its Genesis would be our Armageddon.

In the standard model of particle physics, the "rim" is at a value about 100 million times larger than the measured value of the Higgs field.

The actual form of this new universe would depend on what the new value of the Higgs field turned out to be. In our analogy, this would depend how long the ball can roll down the side of the volcano without encountering an obstacle. This is something that we can't determine in the standard model, because the field values get so strong that we run into quantum gravity, and no one currently knows how quantum gravity works.

Before you start worrying too much, I should note that the calculation of this "rim" for the standard model Higgs involves an extrapolation far, far beyond what we have experimentally probed, so we don't know that the standard model prediction for the Higgs potential energy is correct. If the standard model is just an approximation to a deeper theory, then all bets are off. Even if the standard model prediction was correct, the average time it would take for this process to happen is much longer than the current age of the universe, so it is extremely unlikely to happen anytime soon.

The estimated time for the instability in the standard model of particle physics ranges from 10^{100} to 10^{200} years, while the current age of the universe is about 1.4×10^{10} years.

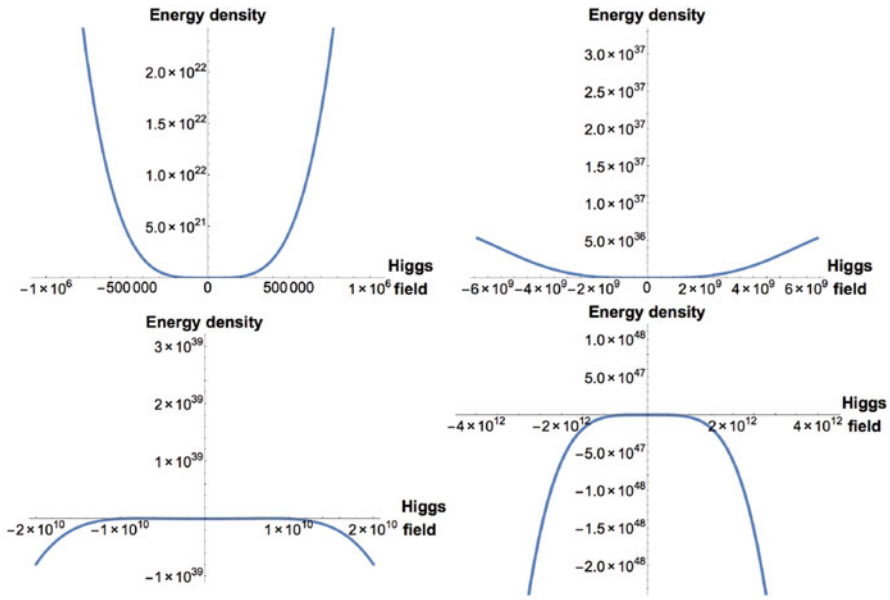


Fig. 9.3 The stored energy plotted as a function of the value of the Higgs field (in GeV or 10^9 electron-volts) over increasingly broader ranges of the Higgs field. As we examine larger values of the Higgs field, the energy stored gets larger as we go to very large values of the Higgs field. However at extremely large values of the Higgs field, around 10^{10} GeV, we see that the energy stored starts to decrease to large, negative values.

9.2 Parallel Universes

▷▷▷ **Misconception !!** The LHC could prove the existence of a parallel universe.

An article from the *Science Times* in 2015 proclaimed: "Large Hadron Collider Could Prove the Existence of Star Trek's Parallel Universe." This was accompanied by pictures of good Spock and evil Spock (you can tell he is evil by the beard; see Fig. 9.4).

Fig. 9.4 The good Spock and the evil Spock from a parallel universe.

In order to understand the (feeble) basis for this headline, we have to know something about the dimensions of space and time. You know that you can specify your position on the surface of the Earth by two numbers, for example, latitude and longitude or even the numbers of two cross streets in a city. This is because the surface of the Earth is two-dimensional. If you actually had a meeting with someone at the intersection of First Street and Fifth Avenue, you might find that you need another dimension; that is, you might have to know what floor of the building your meeting was on. The height above the surface of the Earth is a third dimension. This means that space is three-dimensional: we can specify a position in space by three numbers. Still, to make your meeting, you need to know one more number: you need to know the time of the meeting. This really means that spacetime is four-dimensional. There are lots of other numbers we could use to give information about your meeting like the temperature of the room, or the barometric pressure, but those other numbers do not correspond to dimensions.

To understand a little better why some numbers correspond to dimensions and other numbers do not, it's helpful to think about what happens when we describe positions using two different systems of labeling directions (or coordinates). Suppose that to go from your house to work, you need to go 3 miles east and 4 miles north; that means your office is 5 miles (as the crow flies) from your house. Suppose I had purchased a cut rate GPS system that didn't quite work and that thinks north is in a different direction, then it might tell me that your office is 1 mile "north" and 4.9 miles "east." That would still make your office 5 miles away, but the way I've divided up the distance into "north" and "east" (the coordinates) is different from the usual definition of north and east; see Fig. 9.5.

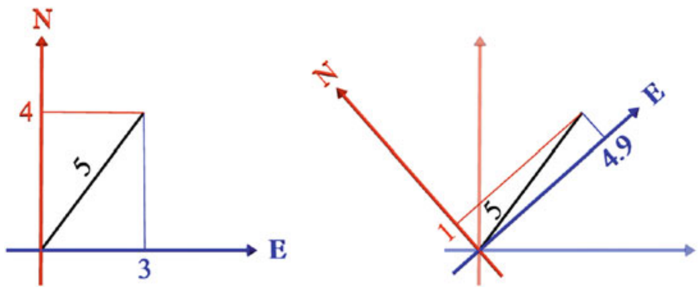


Fig. 9.5 The distance between two points does not depend on which direction we choose to call North, but the part that is North and the part that is East do depend on this choice.

Mathematically we would say that my coordinate system is rotated from your coordinate system, my "North" is a combination of your North and West, and my "East" is a combination of your North and East. Which direction we choose to call North is somewhat arbitrary. We could use the geographic North Pole or the magnetic North Pole, but whichever we use, the distances between two places will be the same. Choosing different coordinate systems is an example of a type of symmetry transformation. In this case, we've chosen two coordinate systems that are related by a rotation which mixes the two dimensions in one coordinate system to come up with two new dimensions in the new coordinate system. This is a general feature of dimensions: under symmetry transformations, dimensions can transform into each other, but the number of dimensions doesn't change. Famously Einstein showed that time is really another type of dimension and that the symmetry transformation that mixes up time and space is just changing our velocity!

A rotation is just one type of a symmetry transformation.

Let's look at a simple example of how space and time could get mixed together. Suppose Anna is holding two flashbulbs. She holds her arms extended sideways and arranges for the two flashbulbs to go off simultaneously. Some of the light from the left flashbulb travels right, toward her nose, and some of the light from the right flashbulb travels left, also toward her nose. If her arms are the same length, then the light from each flashbulb reaches her nose at exactly the same time; see Fig. 9.6.

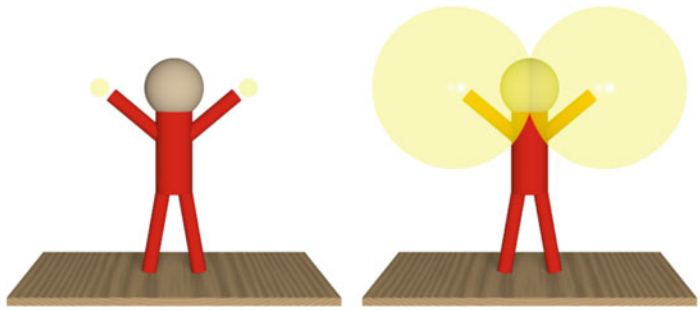


Fig. 9.6 Two lights are flashed on simultaneously and meet at Anna's nose in the middle sometime later.

Now let's imagine what Bob sees as he moves past Anna at some velocity. From Bob's perspective, Anna is moving to the right, and he sees that the light from each flashbulb meets at her nose at exactly the same time. But her nose is moving to the right, toward the right flashbulb and away from the left flashbulb. The light from the right flashbulb travels a shorter distance, so in order to meet at the same time, the light from the trailing flashbulb had to start out earlier; see Fig. 9.7.

Thought experiments like these led Einstein to his theory of special relativity. What Anna thinks of as simultaneous (occurring at one particular time) is, according to Bob, a series of events that occur at different points in time. What Anna thinks of as a space direction (at a single point in time) is a mixture of Bob's time and space. What Bob thinks of as a space direction is a mixture of Anna's time and space directions; see Fig. 9.8.

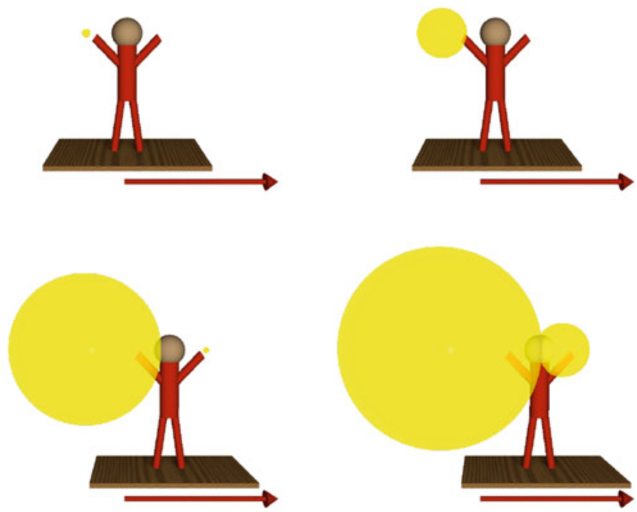


Fig. 9.7 The same two lights are flashed on but viewed by someone moving with respect to Anna. From this point of view, the lights cannot turn on simultaneously if they are both to meet at Anna's nose at the same time, since her nose is moving away from one light and toward the other.

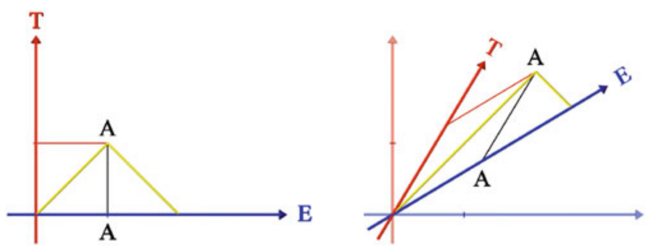


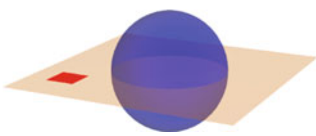
Fig. 9.8 A spacetime "map" of the emission of the simultaneous light flashes. In these maps, light rays travel along *diagonal lines*. The map on the left is from the point of view of Anna, who is moving with the flashbulbs, and the line from A to A represents Anna's nose (where the light flashes eventually meet) moving through time at a fixed position in space. The map on the right shows the point of view of Bob, who sees Anna and the flashbulbs moving toward the East. The flashbulb on the left goes off first, so that the flash can meet up with the flash coming from the right.

▷▷▷ **More Accurate!!** What is simultaneous to you is not simultaneous for someone moving relative to you.

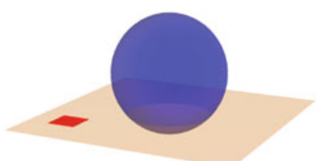
Yes his name was really Edwin Abbott Abbott, with Abbott appearing twice. His father was Edwin Abbott.

Now that we have a better idea of what spacetime dimensions are, we can turn our attention, again, to extra dimensions. The idea of extra dimensions has been popular since the 1800s. In 1884, Edwin Abbott Abbott wrote "Flatland," a very funny story about a square living in a two-dimensional world who tries to understand the idea three dimensions after he is visited by a three-dimensional sphere. At first he dismisses the idea that there could possibly be more than two dimensions, but the sphere lifts him out of Flatland into the third dimension where he has a bird's-eye view of everything, including the insides of his friends. After grasping three dimensions, the square begs to be shown higher dimensions, like the fifth, sixth, and even the tenth dimension. The sphere assures him that it is crazy to imagine that there could possibly be more than three dimensions.

A sphere visiting Flatland looks like a circle to the square.



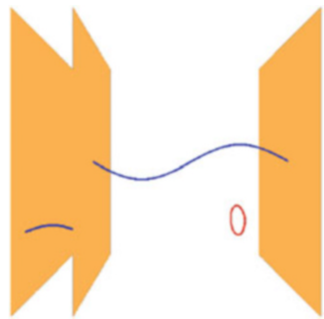
When the sphere moves into the third dimension he appears to shrink to a point and disappear.



In physics the popularity of extra dimensions is primarily due to the rise of string theory, which requires 10 or 11 dimensions in total for self-consistency. It used to be thought that there were six different (and very poorly named) types of string theory: in 10 dimensions, there was Type I, Type IIA, Type IIB, and two types of "heterotic" theories, while in 11 dimensions, there was something called "supergravity." Ed Witten showed that all these different string theories were just approximations to a more complete theory that he called "M theory." For example, if you could take one of the dimensions of the 11-dimensional theory and wrap it around a circle and then imagined making the circle very small, you would end up with one of the 10-dimensional theories. This is also how you could imagine reducing 10-dimensional theories down to the four dimensions of spacetime that we see. You would need six of the dimensions to be very small, for example, being very small circles or perhaps just extending over a small interval.

⊳⊳⊳ **More Accurate !!** String theorists usually prefer that their extra six dimensions preserve some special symmetries, so they use *Calabi-Yau manifolds* rather than circles, spheres, or short intervals.

Jokingly people have referred to "M theory" as the theory formerly known as string, because in addition to the strings, there are variety of extended objects with varying numbers of dimensions. For example, we could consider a membrane which has two dimensions. For brevity this is sometimes referred to as simply a "brane" or a 2-brane to be specific. In this language, a string can be thought of as a 1-brane and a particle as a 0-brane. Now consider a 3-brane, which would be a "brane" that fills up an entire three-dimensional space, like our universe. "M theory" allows for strings to be stuck on this 3-brane. We could even imagine that very short strings on this 3-brane correspond to the particles in our world. Since this theory could have six extra dimensions, we can easily imagine putting another 3-brane, parallel to our universe (the 3-brane we live on), somewhere off in one of the extra dimensions. This parallel 3-brane would be another universe. Depending on what kind of strings are stuck on it, it could have the same laws of physics as our universe or completely different laws of physics.



Parallel 2-branes with strings attached and a closed loop of string moving between them. In "M theory" the particles of gravity, gravitons, correspond to closed loops of string and can move anywhere.

How would we ever see evidence for such a parallel universe? First of all, we would have to establish that there is indeed an extra dimension. One way to do that is to produce particles that are able to leave our 3-brane and travel in the extra dimension. To our limited 3-brane mentality, that would look like particles (with energy) that simply disappeared, just as in Flatland when the sphere is able to vanish by simply moving into the third dimension. It is possible that the Large Hadron Collider could produce such particles. The experimentalists who built the detectors scour their data looking for such events, but so far they have not seen any evidence of an extra dimension. Even if they did find evidence for an extra dimension, this would (by itself) not tell us whether there was an actual parallel universe; so parallel universes remain firmly in the realm of science fiction for now.

Chapter 10

Physics in the News

10.1 Cell Phone Radiation

▷▷▷ **Misconception !!** Scientists raise alarms about cell phones.

A headline from *Mother Jones* magazine from 2015 warns that "Scores of Scientists Raise Alarm About the Long-Term Health Effects of Cell Phones." A photograph shows a woman wincing in pain while holding a cell phone, as an eerie red glow envelops the side of her face next to the phone. Stories about the dangers of cell phones keep popping up year after year. The situation is somewhat similar to the climate change debate. While the vast majority of scientists believe climate change is occurring, there are a few scientists who claim that everyone else is wrong. The media often turns this into a "he said, she said" story, usually without making any comment on the value of the scientific arguments presented.

Unlike politics, science proceeds by examining real evidence, not by presenting "alternative facts."

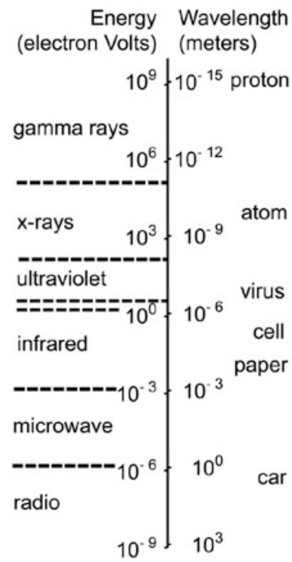
Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_10

One of the problems with the research into the possible health effects of cell phones is that most of the studies are done on very few people for a short amount of time. This means that there is a lot of statistical uncertainty in the results. A further problem is that often the studies are done by choosing people who have some type of brain disease and asking them go back and try to remember some common element that may have caused their disease. As the popularity of cell phones has risen, the likelihood that anybody with a brain cancer also has a cell phone is quite high. This type of study can never tell us if there is a cause and effect relationship or just a random correlation.

On the Web: `http://bit.ly/2n4MOgC`.

The gold standard for medical research is to do a randomized controlled study. In this type of experiment, people are randomly divided into two groups where one is given a treatment (or in this case a cell phone), while the second group does not have access to the treatment (or cell phone). These types of studies are very difficult to do and very expensive, so most of the time they are not done. We are left with a bunch of less than conclusive studies on various topics, and the best we can do is to try to look at the preponderance of evidence. You have probably had the experience of seeing a headline that claims new research shows ______ causes cancer only to find another article a few months later that says new research shows that same item prevents cancer. Just pick one of your favorite foods to fill in the blank, and you will be able to find studies that show that it both causes and protects against cancer. The fact that single (poorly designed) studies can come to either conclusion just demonstrates that we are often trying to measure something that is very subtle and that the noise is bigger than the signal we are looking for. Eventually, with enough studies, one might hope to get an indication of the real situation, if many more studies indicate one or the other outcome. However, it usually takes a very long time, and the more poorly designed the individual studies are, the longer it takes.

In the case of cell phones, however, we can use our knowledge of physics to see if the dire warnings are plausible. Cell phones send and receive radio waves which are a type of electromagnetic radiation. We have lots of different names for different types of electromagnetic radiation. The different names correspond to different wavelengths. For any type of wave, the wavelength is just a distance between two corresponding parts of the wave (see Fig. 7.8). For example, if we look at a water wave, we can measure the distance between two peaks (or two troughs) of the wave, and that gives us the wavelength. Sound waves and electromagnetic waves also have wavelengths; they are just harder to see. We call electromagnetic waves with wavelengths longer than a meter (about a yard) radio waves. We call electromagnetic waves with wavelengths between a meter and a millimeter (about the thickness of a paper clip) microwaves. We call electromagnetic waves with wavelengths up to a thousand times smaller than a millimeter (about the size of the cell) infrared waves. Electromagnetic waves that are few times smaller than that are the waves we can see, visible light. If we go to even shorter wavelengths, we call them ultraviolet light (which have wavelengths about the size of the virus). If they are even smaller, we call them X-rays (which have wavelengths about the size of molecules or atoms). With wavelengths that are even smaller than atoms, we call them gamma rays (Fig. 10.1).



Spectrum of electromagnetic waves, visible light is in the small gap between infrared and ultraviolet.

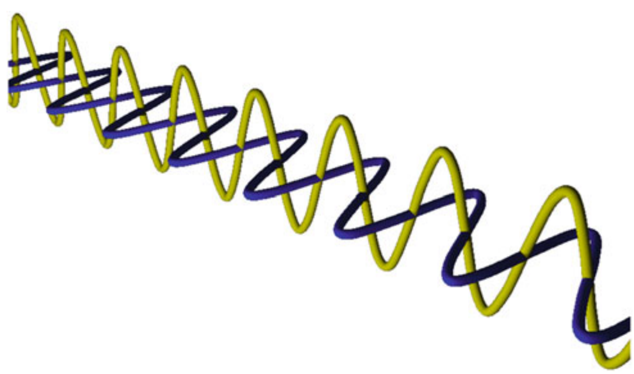


Fig. 10.1 An electromagnetic wave where the electric field is wiggling in the vertical direction and the magnetic field (shown in *dark*) is wiggling in the *horizontal direction*.

Now as we saw earlier, shorter wavelengths correspond to higher energies. Ultraviolet rays can have energies around 10–100 electron-volts, which is enough to scramble your DNA molecules. That is why you want to put on sunblock when you go out in the summer sunshine. While visible light can pass through glass, ultraviolet light is absorbed; that's why you don't get a sunburn when you are indoors even though the sunlight is falling on you through the windows. Microwaves have energies below a thousandth of an electron-volt. That's enough energy to jiggle water molecules around and heat them up (that's how your microwave oven works). Microwaves do not have enough energy to break molecules apart. The wavelengths used for cell phones are longer than those of microwaves and so have even less energy. Nobody has suggested a plausible way for low-power radio waves to damage your brain. Just the infrared radiation in sunlight should be much more dangerous.

You can measure the wavelength used in your microwave oven by removing the rotating platter and placing a large flat piece of chocolate inside and heating it so that it start to melt in a few hot spots. The distance between the hot spots is half the wavelength of the microwaves. The oven sets up a pattern of waves that is not moving in space, the hot spots correspond to points with a peak or a trough (the peaks and troughs get reversed billions of times a second).

One of the more famous studies warning that cell phones are dangerous was a Swedish study that claimed to find an association between cell phones and brain gliomas. The Swedish study was done way back in the 1990s (cell phones were more common in Europe than in the USA at that time). Since then there's been a rapid increase in the number of cell phone users in the USA, so if the study was correct, we would expect to see a rapidly rising rate of gliomas in the USA. However the rate of gliomas is essentially unchanged, see Fig. 10.2.

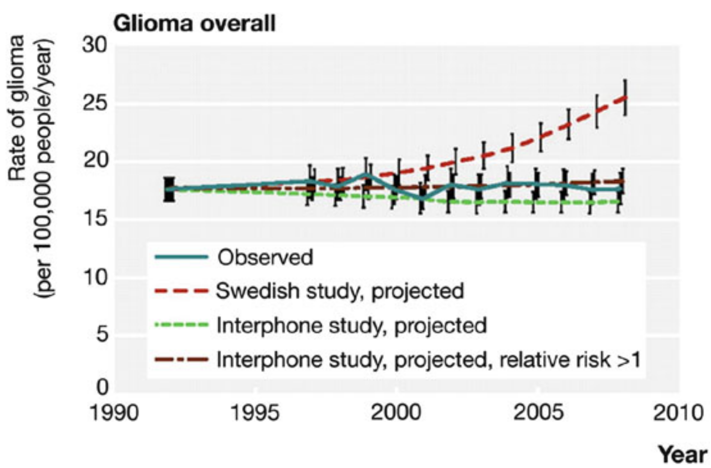


Fig. 10.2 If the Swedish study was correct, the rate of gliomas would have increased dramatically in recent years with increased cell phone use; however, it has remained constant.

So it does not seem plausible that the radio waves emitted by cell phones could cause gliomas, and if they did, we should have seen a rapid growth in the rate of gliomas, which didn't happen. It seems like this is not something worth losing sleep over. Climate change on the other hand . . .

On the Web: `Glioma data:` `bit.ly/gliomarate`.

10.2 Cold Fusion

▷▷▷ **Misconception !!** The cold fusion race just heated up.

Another startling headline (from Wired.co.uk) claimed "The Cold Fusion Race Just Heated Up." Entrepreneur Andrea Rossi claims to have invented a device which when plugged into an electrical outlet produces more energy than it uses. He claims that the device changes one type of metal into another and thus must involve an unknown nuclear process. Even more remarkably, even though these nuclear reactions are occurring (he claims), there is no radiation detected. It all sounds too good to be true and probably is.

Even though nuclear fusion is difficult to control in the lab, we do know that it works and that it is precisely what powers our Sun! Inside of the Sun where the temperature is extremely high and protons are packed extremely densely, two protons will collide often. If one of the protons (through the effects of the W boson of weak interactions) turns into a neutron, a positron, and a neutrino, the proton and neutron can stick together to form deuterium, which is also known as the nucleus of "heavy hydrogen." By adding more and more protons and the conversion of some of the protons to neutrons, stars can build up more complex atoms starting with only hydrogen. These fusion processes can proceed up the periodic table all the way to iron atoms. To go beyond iron atoms takes a supernova.

At the level of quarks, an up quark can turn into a down quark by emitting a W boson. The W boson can then turn into a neutrino and a positron. Changing the up quark to a down quark inside a proton changes the proton to a neutron.

The reason that high temperatures are required is that protons are positively charged, so they electrically repel each other. In order for fusion to happen, we need to get the protons close enough together that they can feel the strong nuclear force. The strong force only operates over a distance of about 10^{-15} m, so we have to get them very close together indeed (basically touching), but it can be done at the center of the Sun. This is also the kind of process that allowed Edwin McMillan and Glenn Seaborg to produce the first transuranic elements (the elements heavier than uranium): neptunium and plutonium. These elements are not normally found on Earth, because they naturally decay, and their lifetimes are much shorter than the age of the Earth. In their experiments McMillan and Seaborg fired deuterium (a proton and a neutron bound together by the strong force) at a uranium target. The deuterium in the beam had a high enough energy that a proton was able to get close enough to stick to the uranium nuclei and form a new element— neptunium.

You may have recognized that McMillan and Seaborg named the new elements after the objects in orbits beyond Uranus, the planet Neptune and the dwarf (former) planet Pluto. The next transuranic element, americium, is used in essentially all home smoke detectors.

The cyclotron that McMillan and Seaborg used was moved from UC Berkeley to UC Davis in 1964 and upgraded. It is still being used, a short walk from my current office.

To perform this feat, McMillan and Seaborg needed a cyclotron to accelerate the deuterium to an energy of 16 million electron-volts. This is because both the deuterium and the uranium nucleus have positive charges and thus repel each other. The deuterium had to have a large velocity (and hence a large energy) in order to get close to the uranium nucleus, see Fig. 10.3.

To put that into perspective, the typical energy of molecules at room temperature is about 0.025 electron-volts. If we heat something up so that it is red hot (about 2600 °F or 1400 °C), the typical energy is 0.15 electron-volts. In the core of the Sun where fusion occurs, the temperature is about 27,000,000 °F (or 15,000,000 °C) corresponding to a typical energy of 9000 electron-volts. Since we know Rossi's machine does not get hot enough to melt, the typical energies can, at most, only be around 1 electron-volt. So it seems extremely unlikely that his machine could actually produce fusion. Of course Rossi would claim that this is some new type of "cold fusion." You probably remember that in 1989 Pons and Fleischman claimed they had produced "cold fusion" but that none of the major labs was able to replicate the results. It is widely thought that what they found was probably some exotic type of surface chemistry involving hydrogen and palladium that produces heat.

On the Web: `http://bit.ly/2nxEzYg`.

Fig. 10.3 A diagram showing a slice of the electromagnetic field lines surrounding two positively charged particles that are approaching each other. The *field lines* bend away from each other because both particles have positive charge. The *top two diagrams* show a low velocity case, while the *bottom diagrams* show a higher velocity. On the *top left*, the particles are approaching each other. On the *top right*, the particles have already reached the point of closest approach and are starting to move apart. They do not get very close because they were approaching at a low velocity. On the *bottom left*, the particles are approaching each other more quickly. On the *bottom right*, the particles are starting to move apart; they get closer than in the previous example because they were approaching at a higher velocity. The higher density of *field lines* between them shows that the repulsive electric field is stronger.

The promise of free energy (or a free lunch) is too tempting for people to ignore, and so people still invest money trying to find a get-rich-quick, energy-for-nothing scheme. Rossi actually licensed his technology to an American firm for a large fee, and this company was supposed to evaluate the technology for one year. At the end of the year, the company said it was no longer interested, and Rossi immediately sued them for breach of contract and for trying to steal his invention. You won't be surprised to learn that Rossi has been convicted of fraud before. What is surprising is that there are still people excited about investing in his company.

Chapter 11

Epilogue

11.1 Spinach: A Cautionary Tale

In most of the misconceptions I've discussed I've been giving the media a hard time for not being careful enough in science reporting. To be fair, cutting edge science is not always easy to understand, and even the experts can disagree at times. Mistakes are inevitable; we can only hope that they are corrected as quickly as possible.

I have one more story to show how difficult it can be. A recent headline in *The Daily Mail* said "Sorry Popeye, spinach doesn't make your muscles big: expert reveals sailor's love of the food was due to a misplaced decimal point." The story goes on to state that a German chemist was measuring the iron content of spinach in the early 1900s and misplaced a decimal point, resulting in the claim that 100 grams of spinach gives you 35 mg of iron. The story goes on to claim that the author of Popeye picked up on this association of spinach and iron and used it to provide Popeye's super strength. This "news" story was taken from a book called "The half-life of facts: why everything we know has an expiration date" by Samuel Arbesman. The book is an examination of how things that we take to be true can turn out later on to be wrong. Unfortunately, the case of the misplaced decimal point in spinach itself was com-

On the Web: `http://dailym.ai/2e1tE66`.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_11

pletely made up in 1972 for an article warning its scientific readers to always check the original sources when doing research. The story was repeated many times, and eventually made its way into Arbesman's book which warned of the various ways that "facts" can turn out to be unreliable. To his credit, Arbesman cleared up the confusion in the second edition of the book. This just goes to show that even well-meaning, intelligent people who are trying to get their stories straight can still make mistakes when trying to present science to the public.

11.2 Moving Forward

There is no danger that the public will run out of misconceptions about science. The real danger seems to be that when the public is confronted by so many conflicting stories about science that it will begin to seem like politics where there no longer seems to be any concept of truth only a competition to see who can put the best spin on any story. Science does not work like politics. There can be controversies, and opposing groups battling for their interpretation to prevail, but in the end, scientists have to back up their claims with evidence, make predictions for new experimental tests, and other scientists have to test those predictions. If the predictions fail, then we have made some progress; we can throw out one more promising-but-wrong idea. Eventually we can decide which of two competing theories best agrees with the facts.

This doesn't mean that one of the theories is correct, just that it is a better approximation to reality than its competitor. Newton's theory of gravity was *the* theory of gravity until Einstein's theory of gravity came along. Both theories agreed in most predictions for the measurements that had been made up to Einstein's time, but they disagreed about how light would bend around the sun. When this was actually measured, Einstein's prediction agreed with the data and Newton's didn't. Newton's theory of gravity was a good approximation when gravity was weak enough, but it was not a good approximation where gravity got strong near the Sun. Someday it is possible that we will test short distance gravity at the quantum level and we will find that we need a new theory that mostly agrees with Einstein's theory except at very short distances. That is how science progresses, a series of better and better approximations.

When one of our approximations fails, it is one of the best times to be a scientist, because it means the time is right to find an even better approximation for describing the world. This is just what people working on the Large Hadron Collider are trying to find: some piece of data that doesn't agree with the standard model of particle physics. If that happens, there will be a worldwide celebration (among particle physicists at least).

Perhaps the biggest misconception about science is that science is a list of facts and that a scientist's job is to collect more facts. Science is really a process for finding better approximations for describing the real world, the facts are what we use to check which is the best approximation so far, and each new approximation leads us to a better understanding of how the Universe really works.

Chapter 12

Extra Material: The Equations Behind the Words

▷▷▷ **Math Alert!** ***double trouble!!*** Here are some of the nuts and bolts that go into the actual calculations in the standard model of particle physics.

12.1 Units and Coordinates

It is usually more convenient to use metric units (technically the International System of Units or SI). A brief list of some of the more useful units is given in Table 12.1. A meter is about a yard long, a liter is about four cups, and a raisin weighs about a gram. Water boils at 373 Kelvin, room temperature is about 293 Kelvin, water freezes at 273° Kelvin, and the remnants of the Big Bang explosion are now at about 3 Kelvin. Nothing can be colder than absolute zero, which is 0 Kelvin.

One of the advantages of using SI is that when going from large to small (or vice versa), it is easier to change to a more convenient unit, since larger units are just multiples of 10 times the basic unit; see Table 12.2. For example, 1 kilometer is 1000 meters and 1 meter is 100 centimeters, so a kilometer is 100,000 centimeters and a kilogram is 100,000 centigrams. Now, quickly, a mile is how many inches?

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_12

Table 12.1 Useful units for different quantities.

Quantity	Unit	Abbrev.
Length	meter	m
Time	second	s
Mass	gram	g
Temperature	Kelvin	K
Frequency	Hertz = 1/s	Hz
Energy	Joules	J
Energy	electron-Volt	eV
Cross section or area	barns	b

Table 12.2 Prefixes and corresponding powers of 10 for SI units.

Power of 10	Prefix	Abbrev.	Number
10^{18}	exa	E	1,000,000,000,000,000,000
10^{15}	peta	P	1,000,000,000,000,000
10^{12}	tera	T	1,000,000,000,000
10^{9}	giga	G	1,000,000,000
10^{6}	mega	M	1,000,000
10^{3}	kilo	k	1,000
10^{2}	hecto	h	100
10^{1}	deca	da	10
10^{-1}	deci	d	0.1
10^{-2}	centi	c	0.01
10^{-3}	milli	m	0.001
10^{-6}	micro	μ	0.000001
10^{-9}	nano	n	0.000000001
10^{-12}	pico	p	0.000000000001
10^{-15}	femto	f	0.000000000000001
10^{-18}	atto	a	0.000000000000000001

Of course, some units can be expressed in terms of more basic units. We know that a velocity—being distance traveled divided by the time taken—is measured in meters per second (m/s), and an acceleration being the rate of change of velocity is measured in m/s^2. Furthermore Newton taught us that force equals mass times acceleration, $F = ma$, so a force is measured in kg m/s^2, and energy can be thought of in terms of work, which is the amount of force exerted through a distance, so in fact, one Joule is equal to 1 kg m^2/s^2. We could also have arrived at the units of energy from Einstein's famous equation that says energy equals mass times the speed of light squared: $E = m\,c^2$. Keeping track of the dimensions of various quantities in an equation is a good way of avoiding mistakes. Sometimes just knowing the dimensions of some quantities allows you to estimate the value of another unknown quantity (without knowing the exact formula that relates them) just by arranging things, so the dimensions match up correctly. This trick, which goes by the fancy name of *dimensional analysis*, was another of James Clerk Maxwell's innovations. While technically eV is not an SI unit, it is widely used in particle physics since it is much more convenient to use for the small amount of energy a single particle typically has. One eV is equal to 1.602×10^{-19} Joules. Barns are not an independent unit, since they are just a unit of area and can be expressed in square meters, 1 b = $10^{-28}\,\mathrm{m}^2$. Barns were introduced when people were studying nuclei by firing protons at them. In order to quantify how likely the proton was to scatter off of (interact with) a given nucleus, an effective area (or *cross section*) was introduced. The typical size of these *cross sections* was $10^{-28}\,\mathrm{m}^2$; hence things were a lot simpler if a smaller unit of area was used. The name started out as a joke, as in "you couldn't hit the broad side of a barn." As we will see, the interesting cross sections at the Large Hadron Collider are much smaller, typically measured in femtobarns, 1 fb = $10^{-43}\,\mathrm{m}^2$ (Table 12.3).

Table 12.3 Greek letters.

Lowercase	Uppercase	Spelling	My pronunciation	Alt. pronunciation
α	A	alpha	"al-fuh "	
β	B	beta	"bay-tah"	"bee-tuh"
γ	$\varGamma$	gamma	"gam-uh"	
δ	$\varDelta$	delta	"dell-tuh"	
ϵ	E	epsilon	"ep-sih-lawn"	"ep-sigh-lun
ζ	Z	zeta	"zay-tah"	"zee-tah"
η	H	eta	"ay-tah"	"ee-tah"
θ	$\varTheta$	theta	"thay-tah"	
ι	I	iota	"eye-oh-tah"	
κ	K	kappa	"cap-pah "	
λ	$\varLambda$	lambda	"lam-duh"	
μ	M	mu	"mew"	
ν	N	nu	"new"	
o	O	omicron	"oh-mih-kron"	
ξ	$\varXi$	xi	"ekszigh"	"ekszee," "cascade"
π	$\varPi$	pi	"pie"	
ρ	P	rho	"roh"	
σ	$\varSigma$	sigma	"sig-muh"	
τ	T	tau	rhymes with cow	"taw"
υ	$\varUpsilon$	upsilon	"up-sih-lawn"	"oops-sigh-lun"
ϕ	$\varPhi$	phi	"fie"	"fee "
χ	X	chi	"kigh"	
ψ	$\varPsi$	psi	"sigh"	"psigh"
ω	$\varOmega$	omega	"oh-may-guh"	"oh-mee-guh," "oh-mig-guh"

Table 12.4 Useful physics constants.

G	Newton's gravitational constant	$6.67 \times 10^{-11}\,\mathrm{m}^3/(\mathrm{kg\,s}^2)$
c	Speed of light	$3.0 \times 10^{8}\,\mathrm{m/s}$
h	Planck's constant	$4.14 \times 10^{-15}\,\mathrm{eVs}$
$\hbar$	Reduced Planck's constant	$6.58 \times 10^{-16}\,\mathrm{eVs}$
α	Fine structure constant	0.007297
k_B	Boltzmann's constant	$8.62 \times 10^{-5}\,\mathrm{eV/K}$
m_e	Electron mass	$9.11 \times 10^{-31}\,\mathrm{kg}$
m_e	Electron mass	$0.511\,\mathrm{MeV/c}^2$
m_p	Proton mass	$1.67 \times 10^{-27}\,\mathrm{kg}$
m_p	Proton mass	$938\,\mathrm{MeV/c}^2$
$M_\oplus$	Mass of the Earth	$5.97 \times 10^{24}\,\mathrm{kg}$
$R_\oplus$	Radius of the Earth	$6.37 \times 10^{6}\,\mathrm{m}$
$R_\odot$	Radius of the Sun	$6.96 \times 10^{8}\,\mathrm{m}$

After writing a lot of equations, one begins to run of letters to use for variable names, so physicists freely use Greek letters. Sprinkling Greek letters in your equations adds a certain panache to your research as well. There is a split on how to pronounce the letters: you say "ep-sigh-lun" and I say "ep-sih-lawn"; you say "oops-sigh-lun" and I say "up-sih-lawn," "zee-tah," "zay-tah," "ee-tah," and "ay-tah"; let's call the whole thing off!

From Table 12.4, we see that everyday units like kilograms (or pounds) are too large and clumsy for delicate elementary particles. We can partially get around this by using the corresponding rest energy of the electron, measured in mega electron-volts (MeV), and quoting the mass as this energy divided by the speed of light squared (i.e., using $E = m\,c^2$).

There are several equivalent ways of specifying a point in a four-dimensional spacetime. We can simply give the values of the coordinates along three perpendicular directions (e.g., East–West, North–South, and up-down) as well as the time. Often these coordinates are grouped in a *vector*, so 5 kilometers East, 4 kilometers North, 1 kilometer high, and 2 seconds after the starting time could be written as a vector like this:

$$(2\,\mathrm{s}, 5\,\mathrm{km}, 4\,\mathrm{km}, 1\,\mathrm{km}). \tag{12.1}$$

where we have chosen (arbitrarily) to write the time first. We could also call the spatial components x, y, and z and the time t, so our example would be written as

$$t = 2\,\mathrm{s}, \quad x = 5\,\mathrm{km}, \quad y = 4\,\mathrm{km}, \quad z = 1\,\mathrm{km}. \tag{12.2}$$

We could also represent the coordinates by a single symbol x_μ where the Greek letter μ runs over four different integer values to distinguish the different coordinates. For historical reasons, the spatial coordinates are label by $\mu = 1, 2, 3$ and the time component is often labeled by $\mu = 0$. It might also be convenient to multiply the time coordinate by the speed of light so that all the coordinates are measured in meters. In the example above, we would write

$$x_0 \;=\; c \times 2\,\mathrm{s} = 6 \times 10^8\,\mathrm{m}, \quad x_1 = 5 \times 10^3\,\mathrm{m}, x_2 = 4 \times 10^3\,\mathrm{m}, \quad x_3 = 1 \times 10^3\,\mathrm{m}. \tag{12.3}$$

12.2 Einstein's Special Theory of Relativity

In the example of rotating the map (see Fig. 9.5), we used the fact that the distance between two points doesn't change if we rotate either the Earth or our choice of coordinates. Suppose that we know the position along the East–West road of the house (call it x_{house}) and of the corner, you turn North to go to the office (call it x_{corner}). We can use the Greek letter Δ, to simplify our equations by using it to indicate a difference, so that Δx is the difference between values of x, in this case between x_{corner} and x_{house}. The distance to the East was 3 kilometers and the distance to the North was 4 kilometers. In equations, we can write the distance between the corner and the house as

$$\Delta x = x_{\text{corner}} - x_{\text{house}} = 3 \text{ km}. \tag{12.4}$$

Similarly along the North–South road, the distance between the corner and the office is

$$\Delta y = y_{\text{office}} - y_{\text{corner}} = 4 \text{ km}. \tag{12.5}$$

Thanks to Pythagoras we know that the distance from the house to the office is 5 kilometers, since

$$(5 \text{ km})^2 = (\Delta x)^2 + (\Delta y)^2 = 9 \text{ km}^2 + 16 \text{ km}^2, \tag{12.6}$$

where the superscript ^{2} (or squared) means "multiply something by itself." Since we have multiplied two lengths, the answer has units of an area, square kilometers.

After we rotate the directions that we call North and East, there are new values for the length of the sides of the triangle with one side in the new East direction, one side in the new North direction, and one side going from the house to the office. We can call the distance along the new East direction x' where the superscript $'$ tells us that this is using our new *coordinate system*. In our example we rotated North by 41.5931°, so we have (in terms of the trigonometric functions cosine and sine, cos and sin for short)

$$\begin{aligned} \Delta x' &= \cos(41.5931°)\,\Delta x + \sin(41.5931°)\,\Delta y = 4.89898\,\text{km}, \\ \Delta y' &= -\sin(41.5931°)\,\Delta x + \cos(41.5931°)\,\Delta y = 1\,\text{km}. \end{aligned} \tag{12.7}$$

But the house and the office haven't moved, we only changed what we called North and East, so the distance between the house and the office can't have changed. Indeed we find

$$(5\,\text{km})^2 = (\Delta x')^2 + (\Delta y')^2 = 24\,\text{km}^2 + 1\,\text{km}^2. \tag{12.8}$$

For a rotation by an arbitrary angle, represented by the Greek letter θ, we would find the same result because of the fact, from trigonometry, that $\cos^2(\theta) + \sin^2(\theta) = 1$ for any angle θ.

We can easily extend this discussion to three dimensions of space. In everyday language, we refer to the three spatial directions as forward/backward, left/right, and up/down, but mathematicians and physicists often call these directions x, y, and z and specify spatial positions by the distances along each of these directions measured from some convenient reference point, which is often called the *origin*. So a physicist would locate his house at x_{house}, y_{house}, z_{house}. For rotations in three dimensions, the *invariant* length L is

$$(\Delta L)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2, \tag{12.9}$$

where Δx, Δy, and Δz are the distances in along the three mutually perpendicular directions measured from the point we are rotating around.

Einstein realized that with a fourth dimension, time, there is another type of coordinate change that we need to be careful about, that is, changing velocities. Changing the velocity of our coordinate system is called a *boost*, and just as a rotation leaves the length ΔL in Eq. (12.9) *invariant*, a *boost* leaves the following quantity *invariant*:

$$(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2, \tag{12.10}$$

where Δt is a time interval and c is the speed of light. We could have multiplied this equation by -1, so that the terms with spatial differences contributed positively, and still found an *invariant*. The overall sign doesn't matter, as long as we stick to our choice consistently; what matters is the relative minus sign between the time contribution and the spatial contribution. Historically physicists on the West Coast of the USA (and most particle theorists) used the convention in Eq. (12.10), following a famous textbook by Bjorken and Drell, while physicists on the East Coast of the USA (and most gravity theorists) used the opposite convention, but we still all manage to get along with each other.

Note that the c^2 in the first term in Eq. (12.10) when multiplying $(\Delta t)^2$ makes this term have the same units as the other terms: length squared. Now consider two *events*, call them 1 and 2, that are separated by a distance

$$\Delta x = x_2 - x_1 \tag{12.11}$$

in the x direction and by a time interval

$$\Delta t = t_2 - t_1. \tag{12.12}$$

In other words event 1 is at position x_1 when a clock at that position reads t_1, and event 2 is at x_2 when the clock at that position reads t_2. In this example Δy and Δz are zero since both events take place at the same distance along the y direction and same height in the z direction. So the relativistic *invariant* separation in this example is

$$(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta x)^2 \tag{12.13}$$

If we start moving in the x direction with a velocity v (i.e., we *boost* to new coordinates that are moving with a velocity v in the x direction relative to our old coordinates), then we have new coordinates given by

$$\Delta t' = \gamma(\Delta t - \frac{\mathrm{v}}{c^2}\,\Delta x) \tag{12.14}$$

$$\Delta x' = \gamma(\Delta x - \mathrm{v}\Delta t) \tag{12.15}$$

$$\Delta y' = \Delta y \tag{12.16}$$

$$\Delta z' = \Delta z, \tag{12.17}$$

where the *boost* factor, represented by the Greek letter γ, is given by

$$\gamma = \frac{1}{\sqrt{1 - \frac{\mathrm{v}^2}{c^2}}}. \tag{12.18}$$

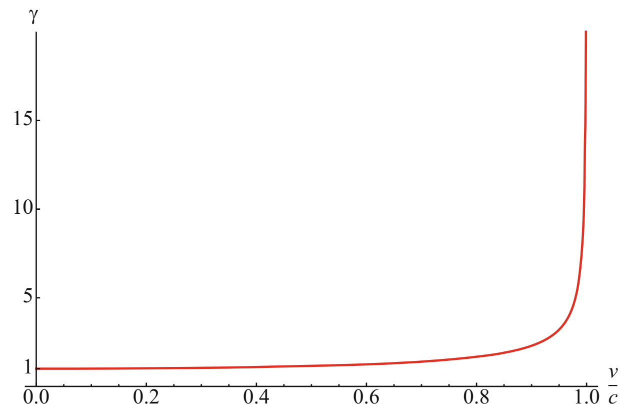


Fig. 12.1 γ stays very close to 1 for velocities much smaller than the speed of light but grows very large as the speed of light is approached.

In the example of Anna and Bob, we can have Anna using the coordinates (t', x', y', z') while Bob uses (t, x, y, z). In the new, moving, coordinate system we can calculate that Δs is given by

$$\begin{aligned}
(\Delta s)^2 &= c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 \qquad (12.19)\\
&= c^2\gamma^2(-\frac{\mathrm{v}}{c^2}\,\Delta x + \Delta t)^2 - \gamma^2(\Delta x - \mathrm{v}\Delta t)^2\\
&= \gamma^2(c^2\Delta t^2 + \frac{\mathrm{v}^2}{c^2}\,\Delta x^2 - 2\mathrm{v}\,\Delta x\Delta t) - \gamma^2(\mathrm{v}^2\Delta t^2 + \Delta x^2 - 2\mathrm{v}\,\Delta x\Delta t)\\
&= \gamma^2((c^2 - \mathrm{v}^2)\Delta t^2 + (\frac{\mathrm{v}^2}{c^2} - 1)\,\Delta x^2)\\
&= \gamma^2(1 - \frac{\mathrm{v}^2}{c^2})(c^2\Delta t^2 - \Delta x^2)\\
&= c^2\Delta t^2 - \Delta x^2, \qquad (12.20)
\end{aligned}$$

which is precisely the same as we found in the old coordinates in Eq. (12.13). If the two events we were considering where the right distance apart so that a photon could be emitted from the first event and arrive at the second event, then the photon would have traveled a distance equal to the speed of light times Δt so

$$\Delta x = c\Delta t \tag{12.21}$$

and also $\Delta s = 0$. In the *boosted* coordinate system, Δs is also zero and if we calculate the speed of light in this *boosted* coordinate system, it is given by distance traveled divided by the time interval:

$$\begin{aligned} \frac{\Delta x'}{\Delta t'} &= \frac{\sqrt{c^2(\Delta t')^2 - (\Delta s)^2}}{\Delta t'} = \frac{\sqrt{c^2(\Delta t')^2}}{\Delta t'} \qquad (12.22) \\ &= \frac{c\Delta t'}{\Delta t'} = c. \qquad (12.23) \end{aligned}$$

This is just what Einstein realized has to be the case: the speed of light measured in one coordinate system and in a *boosted* coordinate system has the same value, c. Or, in other words, the speed of light is independent of how fast we are moving when we measure the speed of light. The price of achieving this agreement with reality (it has been tested experimentally over and over again) is that if we have simultaneous events in one coordinate system (so that $\Delta t = 0$), they are not simultaneous in another *boosted* coordinate system, since from Eq. (12.17) we have

$$\Delta t' = \gamma(-\frac{\mathrm{v}}{c^2}\,\Delta x) \qquad (12.24)$$

The minus sign tells us that, for positive v, the event with the larger value of x happens first in the new coordinate system.

We can also invert these relationships to go from Anna's coordinates to back to Bob's

$$\begin{aligned} \Delta x &= \gamma(\Delta x' + \mathrm{v}\Delta t'), \qquad (12.25) \\ \Delta t &= \gamma(\Delta t' + \frac{\mathrm{v}}{c^2}\,\Delta x') \qquad (12.26) \end{aligned}$$

Notice the change is essentially that the velocity in the formula goes from v to $-$v. This makes perfect sense, since if Bob sees Anna moving at velocity v in a certain direction, then Anna must see Bob moving at the same speed in the opposite direction. The minus sign in the velocity just tells us that v and $-$v represent the same speed but in opposite directions.

In the example of Anna's simultaneously flashing lights discussed earlier, the *boost* velocity is in the Easterly direction, since Bob sees Anna moving to the East, and v is positive. In this case, we find that for Anna's simultaneous events ($\Delta t' = 0$), Bob sees a time difference of

$$\Delta t = \gamma\frac{\mathrm{v}}{c^2}\,\Delta x', \qquad (12.27)$$

and the event with the smaller value of x happens first in the *boosted* coordinates. This is just as we expect: the light leaving from the trailing hand (which covers more distance) meets up at Anna's nose at the same time as the light from the leading hand (which travels for a shorter time).

As the speed v gets closer and closer to the speed of light c, γ gets larger and larger, as shown in Fig. 12.1. If the velocity change is very small compared to the speed of light (as it is in everyday life), then the coordinate change is approximately

$$\begin{aligned} \Delta x' &\approx (\Delta x - \mathrm{v}\Delta t) \qquad (12.28) \\ \Delta t' &\approx \Delta t \quad \Delta y' = \Delta y \quad \Delta z' = \Delta z, \qquad (12.29) \end{aligned}$$

which is indeed what we would naively expect: as the new coordinates move along the Easterly (or x) direction with velocity v, points that were stationary in the old coordinates seem to move to the West with a speed v.

Now consider an object that has length L' along the x direction according to Anna who is moving along with it, how long will it look to Bob? Well, in order to measure its length, Bob will look at the distance between the two ends at some fixed time according to his clock, which means we can use $\Delta t{=}0$ in Eq. (12.15). If we call the length that Bob measures L, then we must have

$$L' = \Delta x' \quad = \quad \gamma(\Delta x - \mathrm{v}\cdot 0) = \gamma L. \tag{12.30}$$

So according to Bob, who sees the object moving at velocity v along the direction of its length, the length is actually

$$L = \frac{L'}{\gamma} \tag{12.31}$$

This is called relativistic length contraction. Objects appear to be contracting when they are moving relative to us. If it had been Bob holding the same object, Anna would have seen it as being shorter.

Similarly a time interval $\Delta t'$ between two events experienced by Anna at the same position in her coordinates will seem to be a longer time interval to Bob, who measures on his clock a time difference between the same two events given by Eq. (12.25):

$$\Delta t \quad = \quad \gamma(\Delta t' + \frac{\mathrm{v}}{c^2}\cdot 0) \tag{12.32}$$

$$\Delta t \quad = \quad \gamma\Delta t'. \tag{12.33}$$

The time that Bob measures between the two events is longer than the time measured by the moving clock, so it seems to him that Anna's clock is running slow.

As we have seen what is simultaneous for Anna is not simultaneous for Bob, and what has a constant position for Anna is moving for Bob, so Bob's positions and times are a mixture of Anna's positions and times. If the position and time of an event measured by Bob are given by x, t, and the corresponding position and time for Anna are x', t', then there is a simple relation between the two sets of observations. The relation is even simpler if Anna agrees to set the zero of her x' and t' coordinates to agree with the zero of Bob's x and t coordinates. In this case the relations is

$$x' = \gamma(x - \mathrm{v}t), \quad t' = \gamma(t - \frac{\mathrm{v}}{c^2}\,x). \tag{12.34}$$

$$x = \gamma(x' + \mathrm{v}t'), \quad t = \gamma(t' + \frac{\mathrm{v}}{c^2}\,x'). \tag{12.35}$$

In Einstein's theory of special relativity when we examine distances and times, the *invariant* quantity under *boosts* is the difference of the squares of time component and the squares of spatial components as seen in Eq. (12.10). Similarly, for energy and momentum, the *invariant* is the square of the energy (E) minus the squares of momentum components (p_x, p_y, p_z) times the speed of light squared:

$$E^2 - c^2(p_x^2 + p_y^2 + p_z^2). \tag{12.36}$$

Now if we are interested in a particular particle that happens to have a mass, m, we can always *boost* to a frame of reference (or coordinate system) where all the components of the momentum are zero (this is the frame of reference where the particle is at rest, aka the *rest frame*). Einstein also told us that the energy of a particle at rest is

$$E_{\text{rest}} = mc^2 \tag{12.37}$$

but this means that we know the value of the *invariant* above, since we know it in one particular frame, it is *invariant* so it must be the same in any frame:

$$E^2 - c^2(p_x^2 + p_y^2 + p_z^2) = m^2c^4. \tag{12.38}$$

Solving for E, we have

$$E = \sqrt{(p_x^2 + p_y^2 + p_z^2)c^2 + m^2c^4}. \tag{12.39}$$

We can always rotate our coordinates so that the x direction is lined up with the direction the particles are moving, so that $p_y = 0$ and $p_z = 0$. Using

$$p = \gamma m\mathrm{v}, \tag{12.40}$$

we can find an even simpler expression for the energy:

$$\begin{aligned} E &= \sqrt{p^2c^2 + m^2c^4} = \sqrt{\frac{m^2v^2c^2}{1 - \mathrm{v}^2/c^2} + m^2c^4} \\ &= \sqrt{\frac{m^2v^2c^2 + (1 - \mathrm{v}^2/c^2)m^2c^4}{1 - \mathrm{v}^2/c^2}} \\ &= \sqrt{\frac{m^2c^4}{1 - \mathrm{v}^2/c^2}} = \gamma\, mc^2. \end{aligned} \tag{12.41}$$

The kinetic energy (the energy associated with motion) is just the difference of total energy and the rest energy

$$K = (\gamma - 1)mc^2. \tag{12.42}$$

Using the Taylor series expansion for small $x \ll 1$

$$\frac{1}{\sqrt{1-x}} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \ldots, \tag{12.43}$$

we find that for small velocities:

$$K \approx \frac{1}{2}m\,v^2, \tag{12.44}$$

which is what was always used before Einstein came along. It was a good enough approximation as long as the velocities were much smaller than the speed of light, which is certainly the case in everyday life.

12.3 Quantum Mechanics and Scattering

The frequency, f, of oscillation of a light wave (in oscillations per second, aka Hertz) is related to the wavelength (the distance between neighboring peaks of the wave). Wavelength is usually represented by the Greek letter λ, so the frequency is given by

$$f = \frac{c}{\lambda}, \tag{12.45}$$

where c is the speed of light. If the wavelength is given in meters, then the speed of light should be given in meters per second, so that meters cancel out, and the result has units of $1/\mathrm{s}$, which is equivalent to Hertz. This formula actually works for any wave if c is replaced by the speed of the wave.

When Max Planck was trying to understand the radiation emitted by a hot object, he found that he could only make his formula agree with experiment if he put in an extra "fudge" factor. He called it the "hilfsgröβe," which is German for "help-factor," hence the abbreviation to h. His idea was that the atoms could only exchange a quantized amount of energy, hf, with the electromagnetic field. Einstein took this a step further, when he explained the photoelectric effect by realizing that light was composed of many photons, each of which carried an energy

$$E = h\,f. \tag{12.46}$$

If we combine these two formulae, and eliminate f, we can relate the energy and the wavelength:

$$E = \frac{h\,c}{\lambda}. \tag{12.47}$$

There is also a corresponding relation between the momentum of a photon and the wavelength

$$p = \frac{h}{\lambda}, \tag{12.48}$$

which agrees with the formula for the Energy (12.39) once we remember that a photon has zero mass ($m = 0$).

One often sees Planck's constant written in another form, the reduced Planck's constant, or h-bar:

$$\hbar = \frac{h}{2\pi}. \tag{12.49}$$

This form is quite popular, partly because there can be many different numbers called h, but there is only one called $\hbar$. When you mention $\hbar$ to a physicist, they immediately know you are talking about quantum mechanics. If a calculation doesn't involve $\hbar$ at all, then we call it "classical" in order to indicate that it involves no quantum mechanics.

In the 1920s it occurred to Arthur Holly Compton that if Einstein was right and light was really made of photons with energy $E = hf$, then you might expect to be able to scatter a photon and an electron just like scattering two electrons or even two billiard balls. The scattering of photons and electrons was first measured by Compton in 1923, which earned him the Nobel Prize 4 years later. One can calculate the final energy that the photon has after scattering off an electron by using the conservation of energy and momentum just as one would do for billiard ball scattering except that here the colliding particles are relativistic (one is even massless), and we can't use the nonrelativistic formula for kinetic energy, as in Eq. (12.44), that we would have used for billiard balls. For the photon we start out with an energy

$$E_\gamma = \frac{hc}{\lambda}, \tag{12.50}$$

and the electron starts at rest with energy

$$E_e = m_e c^2, \tag{12.51}$$

where m_e is the mass of an electron.

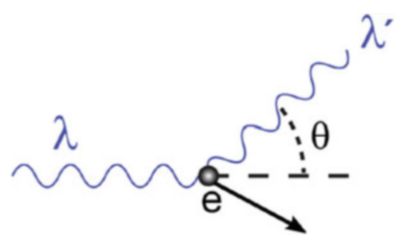


Fig. 12.2 A photon scatters off an electron sitting at rest. After the scattering, the photon goes in a direction at an angle θ away from its original direction.

If the photon emerges from the scattering at an angle θ from its initial direction and with wavelength λ', see Fig. 12.2, then energy and momentum conservation tell us that the shift in the wavelength must be

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta). \tag{12.52}$$

The final photon energy is

$$E' = hf' = \frac{hc}{\lambda'}. \tag{12.53}$$

If the photon goes straight through, which means $\theta = 0$, so $\cos\theta = 1$, there is no change in wavelength, so in this case, the photon's energy does not change. For other angles, the final wavelength is longer, so the photon has lost energy to the electron. When the photon bounces straight back, then $\cos\theta = -1$, and the photon suffers maximum energy loss.

Suppose that we have N_e electrons in our sample and our beam of photons delivers Φ photons per square meter per second onto the target, we call the product of these the *luminosity*:

$$\mathcal{L} = N_e \Phi. \tag{12.54}$$

We can count the number of scattered photons detected per second, R, within a narrow range around a given angle θ, as in Fig. 12.3. This means we count all the photons that pass through a strip. It is conventional to call the area of this strip $d\Omega$. Then in order to get some information about the scattering that is independent of the *luminosity*, we define the *differential cross section*[1] as

$$D(\theta) = \frac{R}{\mathcal{L}\, d\Omega} \tag{12.55}$$

If we integrate overall values of θ and ϕ (the azimuthal angle), we get the total *cross section*:

$$\sigma = \int_0^{2\pi} \int_0^{\pi} D(\theta) \sin\theta d\theta\, d\phi = 2\pi \int_0^{\pi} D(\theta) \sin\theta d\theta \tag{12.56}$$

The total *cross section* has units of area, and one can check that for scattering off of an impenetrable sphere of radius r, the total *cross section* is just πr^2, which is, after all, just the cross-sectional area of a sphere.

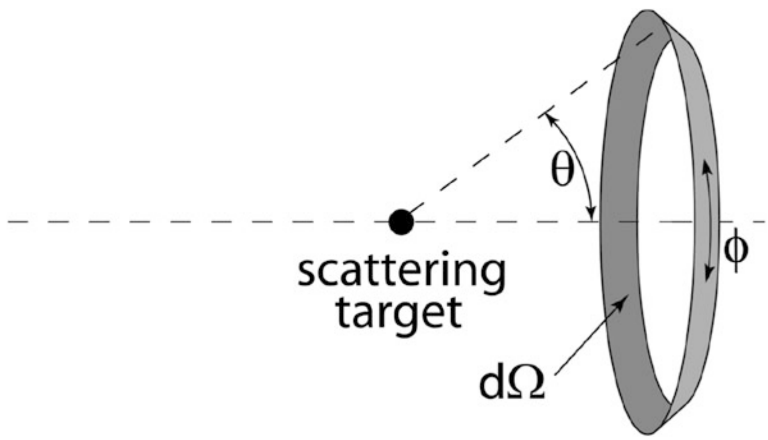


Fig. 12.3 Detecting scattering particles in an area, $d\Omega$, centered at the angle θ. This drawing assumes that the scattering probability is independent of rotations around the horizontal axis through the target, so that we can lump together all the scatterings where the emerging particle intersects the gray strip, that is, we have integrated over the azimuthal angle ϕ.

In 1929, Oskar Klein and Yoshio Nishina were able to use quantum mechanics and special relativity to calculate the differential *cross section* for Compton scattering:

$$D_{\text{Compton}}(\theta) = \frac{1}{2} \left(\frac{\alpha\, \hbar}{m_e\, c} \right)^2 \left(\frac{\lambda}{\lambda'} \right)^2 \left[\frac{\lambda}{\lambda'} + \frac{\lambda'}{\lambda} - 1 + \cos^2\theta \right] \tag{12.57}$$

[1]The differential cross section is often written as $d\sigma/d\Omega$.

where $\alpha \approx 1/137$ is the *fine-structure constant*, which is related to the strength of the interaction between photons and electrons. When the initial photon wavelength is very large, that is, when the photon energy is very small compared to $m_e c^2$, Klein and Nishina's formula reduces to the *cross section* for scattering electromagnetic waves off of an electron found by J.J. Thomson (the discoverer of the electron) in 1906, before quantum mechanics was understood. This means that for long wavelengths the scattering looks classical rather than quantum mechanical. Thomson did not use α in his calculation; he used a classical quantity, Coulomb's constant,[2] which we can write as $k_C = \hbar\, c\, \alpha$. Written in terms of k_C, the Thomson scattering *cross section*

$$D_{\text{Thomson}}(\theta) = \frac{1}{2}\left(\frac{k_C}{m_e\, c^2}\right)^2 \left[1 + \cos^2\theta\right] \tag{12.58}$$

does not go to zero as $\hbar \to 0$, as we expect for a classical result.

Because of the large density of unbound electrons in the Sun, a photon can typically only travel a fraction of a centimeter before it scatters off an electron. This length can be calculated using Thomson's formula (12.58). Individual protons contribute much less to photon scattering, since they are about 2000 times heavier and the mass of the charged particle appears in the denominator of Eq. (12.58), and it is squared. Since the photon can scatter into a completely new direction, it can take a photon 100,000 years to get from the center of the Sun to the surface.

For atoms without too many protons, like carbon, for example, Compton scattering is the dominant scattering process for photons with energies between 10^5 eV and 10^7 eV. For carbon atoms the total *cross section* is about 1 barn, or $10^{-28}\,\text{m}^2$, over this range of energies. This may seem like a very small area, but it is a large *cross section* by particle physics standards. The *cross section* for producing Higgs bosons at the Large Hadron Collider is about 5×10^{-11} b, while many hypothetical particles that are being searched for at the Large Hadron Collider have production *cross sections* around a femtobarn, $10^{-15}\,\text{b} = 1\,\text{fb}$. Searches for dark matter particles have put very stringent upper bounds on their *cross section* for interacting with normal matter. For dark matter particles with masses comparable to the Higgs boson, these *cross sections* must be less than 10^{-21} b!

Calculating *cross sections* can be quite complicated, but for the case of two initial particles, going to two final particle things is fairly simple in the frame of reference where the two initial particles have the same magnitude of their momentum, $p = \gamma m v$, but are traveling toward each other in opposite directions. This is often called the center of mass frame, but a better description is the center of momentum frame: the total momentum in this frame is zero. Often in the actual lab setup, this is not the case, but we can always use special relativity to *boost* our coordinates to this frame where things are simple. To complete our calculation, we need four quantities: the quantum scattering amplitude, $\mathcal{M}$ (which is related to the quantum wavefunctions); the initial momentum of one of the particles, p_i; the final momentum of one of the particles, p_f; and the total center of mass energy

$$E_1 + E_2 = \gamma_1\, m_1\, c^2 + \gamma_2\, m_2\, c^2 \equiv \sqrt{s}, \tag{12.59}$$

[2] Coulomb's constant is often written as $k_C = k_e e^2$, where $e = 1.6 \times 10^{-19}$. Coulombs are the basic unit of electric charge.

where we have introduced the boost *invariant* Mandelstam variable s. Then we can write the differential *cross section* as

$$D_{CM}(\theta) = \frac{1}{64\,\pi^2 s}\,\frac{p_f}{p_i}|\mathcal{M}|^2\,\hbar^2 c^2. \tag{12.60}$$

Generally quantum scattering amplitudes like $\mathcal{M}$ are complex numbers (see Sect. 4.2; we will discuss complex numbers again in Sect. 12.8), but the differential *cross section* is a probability which must be a real number, so it is not surprising that the modulus squared, $|\mathcal{M}|^2$, appears since it is just the sum of the squares of the real and imaginary parts of $\mathcal{M}$.

As an example, consider the process where an electron and positron (e^- and e^+) scatter by annihilating into a muon and anti-muon (μ^- and μ^+) at very high energies where the center of mass energy is much greater than the rest energies of the electrons and muons, $\sqrt{s} \gg m_e c^2, m_\mu c^2$, so that $s \approx 4p_i^2$. The squared amplitude for this process (averaged over spins) is

$$|\mathcal{M}|^2 = 16\,\pi^2\,\alpha^2\,\left(1+\cos^2\theta\right). \tag{12.61}$$

This gives a differential *cross section*

$$D_{CM}(\theta) = \frac{\alpha^2}{4\,s}\left(1+\cos^2\theta\right)\hbar^2 c^2. \tag{12.62}$$

Integrating overall values of the angles θ and ϕ, we get the total *cross section*

$$\sigma_{CM} = 2\pi\int_0^\pi D_{CM}(\theta)\sin\theta d\theta = \frac{4\pi\,\alpha^2}{3\,s}\,\hbar^2 c^2. \tag{12.63}$$

A similar calculation gives the typical lifetime of an unstable particle. Since each particle decays independently of the others, the total number of particles follows an exponential decay law. If we start with N_0 unstable particles at time $t = 0$ then later at time t, we have

$$N(t) = N_0 \exp^{-\varGamma t}, \tag{12.64}$$

where $\varGamma$ is the rate of decay (decays per second) when we factor out the total number of particles. When the particle decays to just two particles, things simplify, and in the frame where the decaying particle is motionless, the calculation is even simpler. If we know the quantum decay amplitude $\mathcal{M}_{\text{decay}}$, the momentum of one of the final particles (p_f), and the mass of the decaying particle, m, then the decay rate is given by

$$\varGamma = \frac{p_f}{32\,\pi^2\,m^2}\int_0^{2\pi}\int_0^\pi |\mathcal{M}_{\text{decay}}|^2\sin\theta d\theta\,d\phi. \tag{12.65}$$

We will use this to calculate the Higgs lifetime in Sect. 12.10.

12.4 Gravity

Newton's universal law of gravitation states that the force between two masses, m_1 and m_2, that are a distance r apart is given by

$$F = -\frac{G\,m_1 m_2}{r^2}, \tag{12.66}$$

where G is Newton's gravitational constant and the minus sign is to remind us that the force acts to pull the masses closer together. Near the surface of the Earth, the gravitational force of the Earth on a mass m is very nearly just a constant multiplying m:

$$F = -\frac{G\,M_\oplus m}{R_\oplus^2} = -mg, \tag{12.67}$$

where (using the values in Table 12.4) we find

$$g = \frac{GM_\oplus}{R_\oplus^2} = 9.8\,\mathrm{m/s^2}. \tag{12.68}$$

The minus sign in (12.67) is telling us that the force is down. So if we drop something near the surface of the Earth, it is accelerated down with the velocity changing at a rate of $9.8\,\mathrm{m/s^2}$. If we lift a mass m a distance L above the surface of the Earth, the amount of energy we put in, aka the work done, is the product of the force and the distance, which is $m\,g\,L$. We call this amount of energy the gravitational potential energy; it is the amount of energy "stored" in the interaction of the mass and the Earth. If we drop the mass, it accelerates down, converting potential energy into kinetic energy (the energy of motion). We can also calculate the gravitational potential energy if we are not close to the surface of the Earth (although it involves calculus since the force changes as we move farther away):

$$V_G = -\frac{G\,m_1 m_2}{r}. \tag{12.69}$$

You can check that the force (12.66) is just -1 times the slope of the potential energy. In other words the force acts in the direction that lowers the potential energy, and the steeper the potential energy curve, the stronger the force.

The gravitational potential energy is very similar to the potential energy between two electric charges, q_1 and q_2, which is

$$V_E = -\frac{k_C\,q_1 q_2}{r}, \tag{12.70}$$

where k_C is Coulomb's constant, named after the French physicist Charles-Augustin de Coulomb (1736–1806) who empirically determined the form of the potential and the associated force. If the charges are opposite (one positive and one negative), then the electric force between them is attractive, and the potential energy grows as they are pulled apart just like Eq. (12.69).

Given the gravitational potential energy (12.69), we can calculate the escape velocity of the Earth. Imagine hitting a baseball hard enough that it keeps going forever. As it goes higher, it slows down as some of the kinetic energy is converted to gravitational potential energy. We describe this by saying that the baseball is losing kinetic energy as it leaves the gravitational potential "well" of the Earth. If we want the baseball to only completely slow down when it is infinitely far away, then we need to start out with just enough kinetic energy so that when it is all converted to gravitational potential energy, we have attained the maximum potential energy, $V_G = 0$ at $r = \infty$. This is how we can calculate the *escape velocity*, which is the minimum velocity that allows the complete escape from Earth's gravity; if we started with a higher velocity, the ball would still be moving when $V_G = 0$. Thus, neglecting air resistance, the escape velocity, $v_{\rm esc}$, is determined by

$$\frac{1}{2}mv_{\rm esc}^2 = \frac{G\,mM_\oplus}{R_\oplus} = m\,g\,R_\oplus \tag{12.71}$$

$$v_{\rm esc} = \sqrt{\frac{2\,G\,M_\oplus}{R_\oplus}} = \sqrt{2\,g\,R_\oplus} = 11.2\,{\rm km/s}. \tag{12.72}$$

Einstein's general theory of relativity explains gravity in terms of the curvature of spacetime. Going back to the spacetime interval given in Eq. (12.10), we first need to consider the case of infinitesimally small intervals, so that we can handle the case where spacetime is continuously changing. As in calculus we denote this infinitesimal limit by replacing Δ by d, so we have

$$ds^2 = c^2dt^2 - [dx^2 + dy^2 + dz^2]. \tag{12.73}$$

We can also use spherical coordinates (r, θ, ϕ, see Fig. 12.4) where

$$x = r\,\sin\theta\cos\phi, \quad y = r\,\sin\theta\sin\phi, \quad z = r\,\cos\theta, \tag{12.74}$$

$$r = \sqrt{x^2 + y^2 + z^2}. \tag{12.75}$$

Because r is *invariant* under rotations, spherical coordinates are very useful when there is spherical symmetry in the problem.

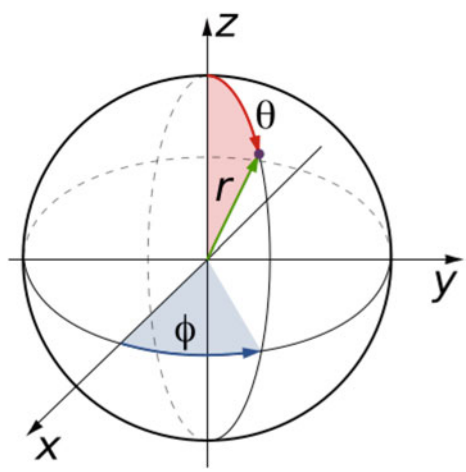


Fig. 12.4 Relation between spherical coordinates (r, θ, ϕ) and Cartesian coordinates (x, y, z). θ is the angle from the z direction, ϕ is the angle (in the x-y plane) from the x direction.

If we had chosen spherical coordinates, then we would write the spacetime interval as

$$ds^2 = c^2 dt^2 - \left[dr^2 + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right]. \tag{12.76}$$

In Einstein's theory, spacetime can stretch and bend, and we account for this by writing the spacetime interval as

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \tag{12.77}$$

where the spacetime metric $g_{\mu\nu}$ is a function of position and μ and ν take the values 0, 1, 2, and 3, so that x^μ represents the four coordinates of spacetime (see the discussion preceding Eq. (12.3)). In Einstein's notation, the repeated indices, μ and ν, in (12.77) are summed over all possible values, so in general there are 16 terms on the right-hand side of the equation.

Einstein's equations relate the metric $g_{\mu\nu}$ to the distribution of matter and energy, and the metric determines how particles move in spacetime. To see how this works in a simple case, we can look at the metric near the surface of the Earth. Since gravity is quite weak near the Earth, Einstein's equations lead to a simple expression (for $r > R_\oplus$) in spherical coordinates:

$$ds^2 = \left(1 - \frac{2GM_\oplus}{r\,c^2}\right)c^2dt^2 - \left(1 + \frac{2GM_\oplus}{r\,c^2}\right)dr^2 - r^2\left(d\theta^2 + sin^2\theta\, d\phi^2\right) \tag{12.78}$$

A particle falling near the Earth just follows the shortest path through the curved spacetime in Einstein's theory. The shortest path on a curved space is called a geodesic. A geodesic on the surface of the Earth is a great circle like the equator or a line of longitude. When you fly between Europe and the USA, you typically fly over Greenland, because the plane is following a geodesic along the Earth's surface. The equation for a geodesic in the spacetime given by Eq. (12.78) can be written as

$$\frac{d^2r}{dt^2} = \frac{1}{2}\frac{d}{dr}g_{00} \tag{12.79}$$

where g_{00} is the coefficient of c^2dt^2 in Eq. (12.78). We can recognize d^2r/dt^2 as the acceleration, a, in the radial (r) direction, so multiplying by the mass of the particle, m, we find

$$ma = -\frac{G\,mM_\oplus}{r^2} \tag{12.80}$$

which is just what Newton would have said.

What Newton would not have known is that the metric Eq. (12.78) implies that clocks run at different rates at different heights above the Earth and that if we shine a beam of light upward, its color changes as it gets higher. First let's consider a clock that ticks every $\Delta\tau$ seconds. If the clock is not moving, then in calculating the *invariant* interval from Eq. (12.78), we can set $dr = 0$, $d\theta = 0$, and $d\phi = 0$. If the clock is very small compared to the size of the Earth (this is true of all clocks that I know of), then we don't have to worry about the difference between dt and Δt, so we can write

$$(\Delta\tau)^2 = \frac{(\Delta s)^2}{c^2} = \left(1 - \frac{2GM_\oplus}{r\,c^2}\right)(\Delta t_{\rm tick})^2, \tag{12.81}$$

this relates interval between ticks, $\Delta\tau$, to the time interval in our coordinate system, $\Delta t_{\rm tick}$. The interesting point is that this depends on r, which is how far we are from the center of the Earth. If we have a clock on the ground and a clock in a tower and send signals between them, then we can see that the clocks run at different rates.

To send a signal, we can shine a light from the bottom clock to the top clock. If the light has a frequency f, then it takes $1/f$ seconds between each crest of the light wave being emitted. We can relate this time interval to the coordinate time interval Δt by using the metric again:

$$1/f = \left(1 - \frac{2GM_\oplus}{R_\oplus\, c^2}\right)^{1/2} \Delta t \tag{12.82}$$

It takes some coordinate time, t, for the first light crest to reach the top of the tower, but it should take the same amount of coordinate time for the next crest, because the metric is not changing with time. So when the light is received at the top of the tower, the coordinate time interval is still Δt s between each crest. If the tower has a height H, then we should observe a time interval

$$1/f_{\rm top} = \left(1 - \frac{2GM_\oplus}{(R_\oplus + H)\, c^2}\right)^{1/2} \Delta t. \tag{12.83}$$

Inverting Eq. (12.82) to find Δt and substituting it, we find

$$\frac{f_{\rm top}}{f} = \left(\frac{1 - \frac{2GM_\oplus}{R_\oplus\, c^2}}{1 - \frac{2GM_\oplus}{(R_\oplus + H)\, c^2}}\right)^{1/2} < 1. \tag{12.84}$$

In other words the frequency of the light has been reduced; this is often referred to as a gravitational redshift, since visible light would be shifted to the longer, red wavelength. Since the photons start out carrying an energy $E = hf$, we can also say that the photons are losing energy as they climb out of the potential well. It also shows us that if we sent out one light pulse for each tick of the bottom clock, we will find that for a given number of ticks of the bottom clock, the top clock has ticked more times. So the top clock runs faster; there are more ticks at the top per tick of the bottom clock.

We can check our answer using the Einstein's equivalence principle, which says that for short distances and times, a gravitational field is just the same as being inside an accelerating elevator. By short, Einstein meant short enough that the gravitational field is approximately constant. Using this idea, we can examine two equivalent situations: we are floating in space with no gravitational field around, and the tower accelerates fast us at $9.8\,\mathrm{m/s^2}$. Alternatively we could be floating in a free-falling elevator, there is a gravitational field, but we don't feel it. Both of these scenarios should give the same results as we found by examining the metric describing the Earth's gravitational field. For example, in the free-falling elevator scenario if we compare our clock with the clock at the top of the tower just as we pass it, we see that it is moving with respect to us, and so by our special relativity calculation, Eq. (12.33), we know that the tower clock is running slow with respect to our clock. If we also check the ground clock (just before we splat), then we will see it is also running slow, but since we have been accelerating while we were falling, we see it as moving faster with respect to us than the tower clock was moving, so it must be running slower than the tower clock as well.

12.5 Black Holes

Einstein's equations allow one to solve for the exact metric around a spherically symmetric mass, M. The solution is called the Schwarzschild metric:

$$ds^2 = \left(1 - \frac{r_s}{r}\right) c^2 dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 \left(d\theta^2 + \sin^2\theta d\phi^2\right), \tag{12.85}$$

where the Schwarzschild radius is given by

$$r_s = \frac{2\,G\,M}{c^2}. \tag{12.86}$$

If the actual radius of the object is less than r_s, then the object is a black hole. When r_s is much less than the radius, then we can Taylor expand the coefficient of the dr^2; let's call it g_{11}; using $x = r_s/r \ll 1$ and

$$\frac{1}{1-x} = 1 + x + x^2 + \ldots \tag{12.87}$$

to find that to leading order, we just reproduce Eq. (12.78) once we set $M = M_\oplus$. For the case of the Earth, $r_s = 8.8\,\mathrm{mm}$ (about 0.35 inches), so Eq. (12.78) is a very good approximation in this case.

After coming up with his famous constant, Planck realized that by using basic *dimensional analysis* he could combine the three basic constants of nature (Newton's gravitational constant, the speed of light, and his constant) in one way to form a length scale:

$$\ell_\mathrm{P} = \sqrt{\frac{\hbar G}{c^3}} = 1.62 \times 10^{-35}\,\mathrm{m}, \tag{12.88}$$

which we now call the Planck length. Following Planck we can combine the three basic constants of nature, G, c, and $\hbar$, to form a mass, the Planck mass scale, and a time scale, the Planck time:

$$M_\mathrm{P} = \frac{\hbar}{\ell_\mathrm{P}\,c} = \sqrt{\frac{\hbar\,c}{G}} = 1.22 \times 10^{19}\,\mathrm{GeV/c^2}, \tag{12.89}$$

$$t_\mathrm{P} = \frac{\ell_\mathrm{P}}{c} = \sqrt{\frac{\hbar\,G}{c^5}} = 5.39 \times 10^{-44}\,\mathrm{s}. \tag{12.90}$$

We can then write r_s in a more suggestive way:

$$r_s = 2\,\ell_{\mathrm{P}}\,\frac{M}{M_{\mathrm{P}}}, \tag{12.91}$$

so the Schwarzschild radius is only large compared to ℓ_{P} when the mass is much larger than M_{P}.

Just as photons leaving the Earth get redshifted, photons leaving the vicinity of a black hole also get redshifted, but since g_{00} goes to zero at $r = r_s$, if they were leaving from the horizon at r_s, then their frequency and energy get redshifted to zero. We can also calculate the escape velocity of a particle of mass m. We need to give it enough energy so that when it gets infinitely far away, it has zero kinetic energy (aka zero momentum), so from Eq. (12.39), its final energy is

$$E = m\,c^2 \tag{12.92}$$

Following a similar logic to our derivation of Eq. (12.84), we can see that at a radius r, the particle must have had an energy

$$E(r) = \frac{m\,c^2}{1 - \frac{r_s}{r}} \tag{12.93}$$

Comparing with Eq. (12.41), we see that the *boost* factor at r was

$$\gamma = \frac{1}{\sqrt{1 - \frac{v_{\mathrm{esc}}^2}{c^2}}} = \frac{1}{1 - \frac{r_s}{r}}, \tag{12.94}$$

so the velocity at r was

$$v_{\mathrm{esc}} = c\,\sqrt{\frac{r_s}{r}} = \sqrt{\frac{2\,G\,M}{r}} \tag{12.95}$$

Which agrees exactly with Eq. (12.72) once we plug in the mass and the radius of the Earth. In the case of a black hole, as we get closer to the horizon, the escape velocity increases. At the horizon the escape velocity becomes equal to the speed of light, c. Since no massive particle can travel at the speed of light and no massive particle can escape once it reaches the horizon, it would need an infinite amount of energy to do so. Even light cannot escape from inside the horizon.

Notice that if we fall through the horizon, then g_{00} and g_{11} both change sign. This means that inside the horizon the coordinate t is a space-like direction, while r becomes time-like. In order to leave the black hole, we need to reverse our r direction, but this means going back in time, which no one knows how to do. As far as we know, if you fell into a black hole, the only way to get out of again is to wait for it to decay, but anything falling in to a large black hole would probably be crushed long before that happens. Since the evaporation proceeds through emitting elementary particles, your particles might come out separately, but they wouldn't really be you anymore.

Using Hawking's evaporation calculation, we can also estimate the approximate lifetime of a black hole of mass m as

$$t \approx t_{\mathrm{P}} \left(\frac{m}{M_{\mathrm{P}}} \right)^3 . \tag{12.96}$$

For a black hole with 30 solar masses ($M = 6 \times 10^{31}$ kg), like the ones LIGO saw, that comes out to about 4×10^{67} years, much longer than the current age of the Universe, which is about 1.4×10^{10} years.

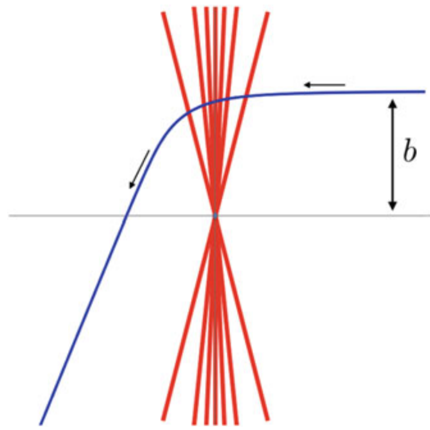


Fig. 12.5 A particle, traveling from *right* to *left*, scattering through the boosted gravitational field of a proton, with the *impact parameter*, b, shown. The motion of the proton is subtracted out so it is at a fixed position in the diagram.

Now let's examine how we could make a black hole during a particle collision. Suppose we had two protons headed toward each other at high speed and that if they both kept going along straight lines, they would miss each other by a distance b. This distance is called the *impact parameter*. Things are a little simpler if we try to catch up with one of the protons, so that from our point of view, it is moving slowly, and we only have to deal with one high-speed proton. From that perspective, the gravitational field lines of the fast proton will bunch up in the sideway direction, as shown in Fig. 12.5. When the proton is moving at almost the speed of light, and has a large momentum $p = \gamma\, m_p\, v \gg m_p\, c$, then the field lines are compressed into a single sheet with the proton at the center. When the slow-moving proton passes through this sheet of gravitational field, it will be swung by the gravitational force so that it is headed quite close to the fast proton. The distance between the two protons gets as small as

$$r_{\min} = \frac{M_{\mathrm{P}}^2\, b^2\, c^2}{8\, p\, \hbar} \tag{12.97}$$

which can be very small (much smaller than b) for a large momentum p. The protons can be so close that we have to take into account their escape velocity, which is

$$v_{\mathrm{esc}} = \sqrt{\frac{2\,G\,m_p}{r_{\mathrm{min}}}} = \frac{4\,\hbar}{M_{\mathrm{P}}^2\,b}\sqrt{\frac{p\,m_p}{c}}. \tag{12.98}$$

If p is large enough, then v_{esc} is greater than the speed of light, and we know that a black hole has formed.

Going back to the point of view of someone standing at CERN with the two protons rushing at each other with the same speeds, but going in opposite direction, we can use the fact that the total center of mass energy of the collision is *invariant* under *boosts* and is given by

$$E_{\mathrm{tot.}}^2 = s \approx 2\,p\,m_p\,c^3 \tag{12.99}$$

so the condition to form a black hole with a fixed $E_{\mathrm{tot.}}$ is that the impact parameter must be sufficiently small:

$$b < b_{\mathrm{crit}} = \frac{\hbar\sqrt{s}}{M_{\mathrm{P}}^2\,c^3} = \frac{G\sqrt{s}}{c^4} \tag{12.100}$$

This corresponds to an effective *cross section* of

$$\sigma \approx \pi b_{\mathrm{crit}}^2 = \frac{\pi\,\hbar^2\,c^2}{M_{\mathrm{P}}^2\,c^4}\,\frac{s}{M_{\mathrm{P}}^2\,c^4} \tag{12.101}$$

Comparing to Eq. (12.63), we see that this *cross section* grows with s instead of falling with s but also that it will not be significant unless $\sqrt{s}$ is comparable to $M_{\mathrm{P}}c^2$ which is not the case for the Large Hadron Collider in a four-dimensional spacetime with Einstein gravity. In the next section, we will see how it could be significant if there are extra dimensions.

12.6 Black Holes in Extra Dimensions

Suppose that there are n extra dimensions that are too small for us to easily resolve. To be more concrete, let us assume that for distances smaller than a certain value, R_*, space has $3+n$ dimensions. For example, if $n = 2$, then the two extra dimensions could be a sphere of radius R_*, so at each point in space, we could move along the directions of the extra sphere. As we saw in Chap. 8, the gravitational field lines would spread out more quickly in $3+n$ dimensions, so for distances $r < R_*$, we would find a gravitational potential given by

$$V_{r<R_*} = -\frac{G_*\,m_1 m_2}{r^{1+n}}. \tag{12.102}$$

For this potential energy to smoothly match on to the ordinary gravitational potential, we must have

$$V_{r=R_*} = -\frac{G_*\,m_1 m_2}{R_*^{1+n}}. = -\frac{G\,m_1 m_2}{R_*} = V_G, \tag{12.103}$$

so we see that

$$G_* = G\, R_*^n. \tag{12.104}$$

In a world of $3+n$ spatial dimensions, G_* is the fundamental Newton's constant that describes the strength of gravity, and G is just the effective strength of gravity that we happen to measure because we are too big and clumsy to resolve the extra dimensions. Since G has units of $\mathrm{m}^3/(\mathrm{kg\,s}^2)$, we see that G_* has units of $\mathrm{m}^{3+n}/(\mathrm{kg\,s}^2)$. This means that we can define a Planck length, a Planck mass, and a Planck time that are appropriate for $3+n$ dimensions:

$$\ell_* = \left(\frac{\hbar G_*}{c^3}\right)^{\frac{1}{n+2}}, \tag{12.105}$$

$$M_* = \frac{\hbar}{\ell_*\, c} = \left(\frac{\hbar^{n+1}}{G_* c^{n-1}}\right)^{\frac{1}{n+2}}, \tag{12.106}$$

$$t_* = \frac{\ell_*}{c} = \left(\frac{\hbar\, G_*}{c^{n+5}}\right)^{\frac{1}{n+2}}. \tag{12.107}$$

Since we haven't seen any black holes produced at the Large Hadron Collider yet, we expect that $M_* \geq 1\,\mathrm{TeV}/c^2$. Using $M_* \approx 1\,\mathrm{TeV}/c^2$, we can get an estimate for R_* from Eq. (12.104). For $n=1$ we get $R_* \approx 3\times 10^13$ m and for $n=2$ we get $R_* \approx 2$ mm, and these are both ruled out by our current tests of gravity as discussed in Sect. 8.1. But for $n=3$, we get $R_* \approx 10^{-8}$ m, which is certainly allowed, and for larger values of n, we get even smaller distance scales, so these cases are perfectly consistent with what we know so far.

We can write the Schwarzschild metric in a $3+n$ dimensional space as:

$$ds^2 = \left(1-\left(\frac{r_{s,n}}{r}\right)^{1+n}\right)c^2dt^2 - \left(1-\left(\frac{r_{s,n}}{r}\right)^{1+n}\right)^{-1} dr^2 - r^2 d\Omega_{2+n}, \tag{12.108}$$

where $d\Omega_{2+n}$ represents an infinitesimal area on a generalized sphere in $3+n$ dimensions with a $2+n$ dimensional surface and

$$r_{s,n} = \left(\frac{2\,G_*\,M}{c^2}\right)^{\frac{1}{1+n}}. \tag{12.109}$$

Now we can redo our calculation of the escape velocity from Sect. 12.5. For a particle of mass m to escape a black hole out to infinity, at a distance r, it must have had an energy

$$E(r) = \frac{m\,c^2}{1-\left(\frac{r_{s,n}}{r}\right)^{1+n}} = \frac{m\,c^2}{\sqrt{1-\frac{v_{\rm esc}^2}{c^2}}} \tag{12.110}$$

so

$$v_{\rm esc} = c\sqrt{\left(\frac{r_{s,n}}{r}\right)^{1+n}} = c\sqrt{\frac{2\,G_*\,M}{r^{1+n}}} \tag{12.111}$$

Repeating our calculation of the black hole production *cross section*, we find

$$b_{\mathrm{crit},n} = \left(\frac{G_* \sqrt{s}}{c^4}\right)^{\frac{1}{1+n}} \tag{12.112}$$

and

$$\sigma_n \approx \pi b_{\mathrm{crit},n}^2 = \frac{\pi \hbar^2 c^2}{M_*^2 c^4} \left(\frac{s}{M_*^2 c^4}\right)^{\frac{1}{1+n}} \tag{12.113}$$

So now we can have a significant production *cross section* at the Large Hadron Collider, where $\sqrt{s} \approx$ few TeV, if M_* is close to 1 TeV/c^2.

With n extra dimensions, black holes can still decay through Hawking radiation. The approximate lifetime of a mini black hole of mass M is

$$t \approx t_* \left(\frac{M}{M_*}\right)^{(n+3)/(n+1)}. \tag{12.114}$$

For $n = 6$ this gives an estimate of 10^{-22} s, about the same as the lifetime of the Higgs boson.

12.7 The Large Hadron Collider

The protons in the Large Hadron Collider beam during the 2016 run had an energy of 6500 GeV, so that the total energy in the head-on collision of two protons was 13000 GeV. Given the proton mass in Table 12.4, we find that the protons had a *boost* factor of

$$\gamma = 6930. \tag{12.115}$$

Using Eq. (12.18), we see that corresponds to the protons traveling with a velocity that is 99.99999896% of the speed of light. The approximate wavelength that these protons can probe is

$$\lambda = \frac{c}{f} = \frac{hc}{6500\,\mathrm{GeV}} = 1.9 \times 10^{-19}\,\mathrm{m}. \tag{12.116}$$

During the 2016 run, the Large Hadron Collider obtained a peak *luminosity* of $1.38 \times 10^{38}/\mathrm{m}^2 s$, or 1.38×10^{10} inverse barns per second. The actual *luminosity* goes up and down as various technical difficulties arise or are overcome. For each period of time with a constant *luminosity*, we can multiply by the length of time that *luminosity* was held, and if we add these all together, we get the total *integrated luminosity*. In the 2016 run, the Large Hadron Collider obtained an *integrated luminosity* of 40 inverse femtobarns. If the *cross section* for producing a new particle was 1 femtobarn, then we could have produced 40 of these particles; obviously having a higher *integrated luminosity* is very beneficial for finding new particles, especially if they have even smaller production *cross sections* than 1 femtobarn. It is hoped that eventually the Large Hadron Collider will collect a total *integrated luminosity* of 300–1000 inverse femtobarns.

12.8 Complex Numbers

Before we get to the Higgs field, we need to remember a few things about complex numbers (which we first heard about in Sect. 4.2), since the value of the Higgs field at any point in space is a complex number. First recall that the "imaginary" number i is the square root of -1, so $i^2 = -1$ (Fig. 12.6). We can write any complex number, z, as either the sum of a real number and i times another real number,

$$z = x + i\,y, \tag{12.117}$$

or as a real magnitude (or modulus), call it r, for example, times a complex phase,

$$z = re^{i\theta}. \tag{12.118}$$

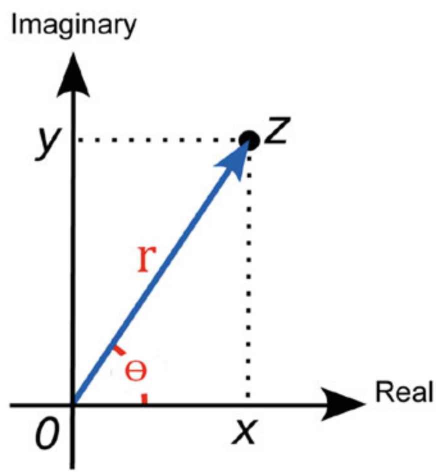


Fig. 12.6 A complex number, z, represented by real and imaginary parts, x and y, and also by modulus r and angle θ.

The two ways of writing the same number are related by thinking about the geometry of the two-dimensional plane of complex numbers, using the real part as the x axis and the imaginary part as the y axis. Using Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta, \tag{12.119}$$

we find

$$x = r\,\cos\theta, \quad y = r\,\sin\theta, \quad |z|^2 = r^2 = x^2 + y^2. \tag{12.120}$$

Whenever we make a measurement, we will always find a real number, so whenever our theory involves a complex number, like z, we often will find that our prediction involves the modulus squared: $|z|^2$.

12.9 Group Theory: Rotations and Spinors

We will also need to know a little about group theory in order to learn a little about the Higgs field. The simplest group that you are probably familiar with is the group of rotations. We can do three types of independent rotations in a three-dimensional space: rotations around the three perpendicular directions or axes. We can call the three axes up, right, and forward, but usually physicists call them x, y, and z. Mathematically rotations form a group because any combination of individual rotations just amounts to one other single rotation around a particular axis (that axis does not have to align with x, y, or z).

We can represent a point in space by a three-component vector, $\vec{v} = (x, y, z)$, where there is one component for each direction, indicating the distance to travel from a reference point along that direction in order to arrive at the specified point. With this choice of coordinates, the reference point is represented by a vector with all zeros: $(0, 0, 0)$. This reference point is also called the *origin*. If we think about very tiny (infinitesimal) rotations around the *origin* (rotations that leave the *origin* untouched), then we can represent the three types of rotations in terms of three matrices:

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{12.121}$$

These three matrices are called the generators of the rotation group. Matrices act on vectors by matrix multiplication: each row is paired up with the vector (written as a column) to produce a new number. Since there are three rows, we get three new numbers, which make up a new vector. For example,

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + jz \end{pmatrix}, \tag{12.122}$$

We can also multiply matrices to get a new matrix of the same size:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} \begin{pmatrix} k & \ell & m \\ n & o & p \\ q & r & s \end{pmatrix} = \begin{pmatrix} (ak + bn + cq) & (a\ell + bo + cr) & (am + bp + cs) \\ (dk + en + fq) & (d\ell + eo + fr) & (dm + ep + fs) \\ (gk + hn + jq) & (g\ell + ho + jr) & (gm + hp + js) \end{pmatrix}. \tag{12.123}$$

Using matrix multiplication, we can work out how the three generators act on a vector:

$$L_x \vec{v} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -z \\ y \end{pmatrix} \tag{12.124}$$

$$L_y \vec{v} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix} \tag{12.125}$$

$$L_z\vec{v} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}. \tag{12.126}$$

Any rotation by an angle θ around an axis pointing along the direction $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$ can be written as an exponential. For convenience we can choose $\vec{\omega}$ to be a unit vector, so that $\omega_x^2 + \omega_y^2 + \omega_z^2 = 1$. Then the rotation matrix is given by

$$R = e^{\theta(\omega_x L_x + \omega_y L_y + \omega_z L_z)}, \tag{12.127}$$

and the rotated vector is given by

$$\vec{v}\,' = R\,\vec{v}. \tag{12.128}$$

It may seem confusing to talk about the exponential of a matrix, but we can make sense of it by writing out the infinite series representation:

$$e^M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \ldots \tag{12.129}$$

where I is the identity matrix with ones down the diagonal and zeroes elsewhere, so that $I\,\vec{v} = \vec{v}$.

Things are a little simpler for infinitesimal rotations, where we take the rotation angle to be infinitesimal, i.e., $\theta = \epsilon$, where $\epsilon \ll 1$. After we do an infinitesimal rotation of the vector $\vec{v}$, the new vector $\vec{v}\,'$ is given approximately by dropping terms as small as ϵ^2 or smaller:

$$\begin{aligned} \vec{v}\,' &\approx [I + \epsilon\,(\omega_x L_x + \omega_y L_y + \omega_z L_z)]\,\vec{v} \\ &\approx \vec{v} + \epsilon\,(\omega_x L_x \vec{v} + \omega_y L_y \vec{v} + \omega_z L_z \vec{v}) \\ &\approx \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \epsilon \left(\omega_x \begin{pmatrix} 0 \\ -z \\ y \end{pmatrix} + \omega_y \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix} + \omega_z \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} \right) \\ &\approx \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \epsilon \begin{pmatrix} \omega_y z - \omega_z y \\ \omega_z x - \omega_x z \\ \omega_x y - \omega_y x \end{pmatrix}. \end{aligned} \tag{12.130}$$

Writing $(\omega_x, \omega_y, \omega_z)$ in terms of spherical angles, as in Eq. (12.74),

$$(\omega_x, \omega_y, \omega_z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{12.131}$$

we find

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \epsilon \begin{pmatrix} z\sin\theta\sin\phi - y\cos\theta \\ x\cos\theta - z\sin\theta\cos\phi \\ y\sin\theta\cos\phi - x\sin\theta\sin\phi \end{pmatrix}. \tag{12.132}$$

This is just an infinitesimal rotation around an arbitrary axis, given by $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$ and by angle ϵ. If we pick $\vec{\omega}$ to point in the z direction (so that $\cos\theta = 1$), we find

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x - \epsilon y \\ y + \epsilon x \\ z \end{pmatrix}. \tag{12.133}$$

which is just the infinitesimal rotation in the two-dimensional x-y plane. Compare this with Eq. (12.7), but replace the angle in degrees by the infinitesimal angle $-\epsilon$ in radians. (Remember that $90° = \pi/2$ radians.) The rotation is by $-\epsilon$ because in Eq. (12.7) we rotated the coordinate system, while now we are rotating objects at particular points, leaving the coordinate system as it was. To make the comparison you need to use the Taylor series expansions

$$\begin{aligned} \cos\epsilon &= 1 - \frac{\epsilon^2}{2} + \ldots \\ \sin\epsilon &= \epsilon - \frac{\epsilon^3}{6} + \ldots \end{aligned} \tag{12.134}$$

We can also define a dot product between two vectors, $\vec{v}_1 = (x_1, y_1, z_1)$ and $\vec{v}_2 = (x_2, y_2, z_2)$, by

$$\vec{v}_1 \cdot \vec{v}_2 = \begin{matrix} (x_1, y_1, z_1) \\ \\ \\ \end{matrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = x_1 x_2 + y_1 y_2 + z_1 z_2. \tag{12.135}$$

You can check that the group of rotations leaves the dot product *invariant* (see Eq. 12.9), just like *boosts* leave spacetime interval (12.10) *invariant*.

Mathematicians have given the group of rotations in three dimensions the name $SO(3)$; it is the group that acts on real, three-component vectors and leaves the dot product *invariant*. A closely related group is $SU(2)$, which acts on complex, two-component vectors. It is common to call any complex, two-component vector by the nickname "spinor," because this is the mathematical object that describes the spin of a spin one half particle like an electron. If we have two such spinors, (u_1, d_1) and (u_2, d_2), then the combination that is left *invariant* by $SU(2)$ transformations is

$$(u_1, d_1) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_2 \\ d_2 \end{pmatrix} = d_1 u_2 - u_1 d_2 \tag{12.136}$$

The generators of the $SU(2)$ transformations are the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{12.137}$$

An $SU(2)$ rotation on the spinor $s = (u, d)$ can be written in terms of a three-component vector $(\alpha_1, \alpha_2, \alpha_3)$ as

$$s' = e^{i(\alpha_1 \sigma^1 + \alpha_2 \sigma^2 + \alpha_3 \sigma^3)} s. \tag{12.138}$$

The combination $d_1 u_2 - u_1 d_2$ is exactly *invariant* under this transformation.

An infinitesimal $SU(2)$ rotation with $\alpha_1, \alpha_2, \alpha_3 \ll 1$ on the spinor $s = (u, d)$ is given by

$$s' \approx \left(1 + i(\alpha_1 \sigma^1 + \alpha_2 \sigma^2 + \alpha_3 \sigma^3)\right) s \tag{12.139}$$

$$\approx \left(1 + i(\alpha_1 \sigma^1 + \alpha_2 \sigma^2 + \alpha_3 \sigma^3)\right) \begin{pmatrix} u \\ d \end{pmatrix} \tag{12.140}$$

$$\approx \begin{pmatrix} u \\ d \end{pmatrix} + i\left(\alpha_1 \begin{pmatrix} d \\ u \end{pmatrix} + i\alpha_2 \begin{pmatrix} -d \\ u \end{pmatrix} + \alpha_3 \begin{pmatrix} u \\ -d \end{pmatrix}\right) \tag{12.141}$$

$$\approx \begin{pmatrix} u \\ d \end{pmatrix} + i \begin{pmatrix} (\alpha_1 - i\alpha_2)d + \alpha_3 u \\ (\alpha_1 + i\alpha_2)u - \alpha_3 d \end{pmatrix}. \tag{12.142}$$

Applying this transformation to the *invariant* in (12.136), we find:

$$\begin{aligned} d'_1 u'_2 - u'_1 d'_2 &\approx (d_1 + i(\alpha_1 + i\alpha_2)u_1 - i\alpha_3 d_1)(u_2 + i(\alpha_1 - i\alpha_2)d_2 + i\alpha_3 u_2) \\ &\quad - (u_1 + i(\alpha_1 - i\alpha_2)d_1 + i\alpha_3 u_1)(d_2 + i(\alpha_1 + i\alpha_2)u_2 - i\alpha_3 d_2) \\ &\approx d_1 u_2 - u_1 d_2 + i(\alpha_1 + i\alpha_2)(u_1 u_2 - u_2 u_1) + i(\alpha_1 - i\alpha_2)(d_1 d_2 - d_2 d_1) \\ &\quad + i\alpha_3(d_1 u_2 - d_1 u_2 + u_1 d_2 - u_1 d_2 -) \\ &\approx d_1 u_2 - u_1 d_2 \end{aligned} \tag{12.143}$$

What Wolfgang Pauli realized when he was trying to understand the spin of the electron is that we can accurately describe quantum spins using this $SU(2)$ group theory. First we introduce the spin operator which is a vector (of matrices) in three-dimensional space:

$$\vec{S} = (S_x, S_y, S_z) = \frac{\hbar}{2}(\sigma^1, \sigma^2, \sigma^3). \tag{12.144}$$

The coefficient $\hbar/2$ just represents the magnitude of one half unit of spin. We can think of a spinor $\chi = (u, d)$ as the wavefunction for spin if we arrange that $|u|^2 + |d|^2 = 1$. Then we can calculate the average value of the spin measured along the z direction by evaluating

$$\langle\chi|S_z|\chi\rangle \equiv (u^*, d^*)\,\frac{\hbar}{2}\sigma^3 \begin{pmatrix} u \\ d \end{pmatrix} = \frac{\hbar}{2}\left(|u|^2 - |d|^2\right). \tag{12.145}$$

So a spinor $\chi_{\text{up}} = (1, 0)$ is always measured to have spin $+\hbar/2$ in the z direction, while a spinor $\chi_{\text{down}} = (0, 1)$ is always measured to have spin $-\hbar/2$ in the z direction. χ_{up} and and χ_{down} are the *eigenstates* (or eigenspinors) of S_z; any other states are superpositions of these two *eigenstates* and do not have a well-defined value for the spin in the z direction; they merely have a certain probability to be measured spin up ($S_z = +\hbar/2$) or spin down ($S_z = -\hbar/2$) as shown in Eq. (12.145). The quantum weirdness gets even stranger when we realize that the *eigenstates* of S_x and S_y are completely different: it is not possible to find a spinor that is an eigenstate of more than one of these three spin components! Also since $\sigma^1\sigma^2$ is not equal to $\sigma^2\sigma^1$, we get a different result depending on the order we act with S_y and S_x. If we act with S_y and then S_x (in our equations the earliest times are to the right), we find:

$$\langle\chi|S_x S_y|\chi\rangle = \frac{\hbar^2}{4}(u^*, d^*)\sigma^1\sigma^2 \begin{pmatrix} u \\ d \end{pmatrix} = i\frac{\hbar^2}{4}(u^*, d^*)\sigma^3 \begin{pmatrix} u \\ d \end{pmatrix} = i\frac{\hbar^2}{4}\left(|u|^2 - |d|^2\right). \tag{12.146}$$

while if we act with S_x and then S_y we find

$$\langle\chi|S_y S_x|\chi\rangle = \frac{\hbar^2}{4}(u^*, d^*)\sigma^2\sigma^1 \begin{pmatrix} u \\ d \end{pmatrix} = -i\frac{\hbar^2}{4}(u^*, d^*)\sigma^3 \begin{pmatrix} u \\ d \end{pmatrix} = i\frac{\hbar^2}{4}\left(|d|^2 - |u|^2\right). \tag{12.147}$$

When things like this happen, we say that the operators are incommensurate or that they don't commute.

12.10 The Higgs

In the standard model of particle physics, the Higgs field transforms as a spinor of the $SU(2)$ group associated with the electroweak interactions. This is a completely new $SU(2)$ group not associated with spin but with two possible values of the electroweak charge. In the early Universe, when the Higgs field vanishes (i.e., the average value of the Higgs field is zero), the electroweak $SU(2)$ group is a perfectly well-behaved symmetry, and curiously left-handed electrons and left-handed electron neutrinos form a spinor as well. Here left-handed means that the spin of the particle is in the opposite direction of its momentum. In the early Universe, electrons, neutrinos, and quarks are all massless, so we can't flip the momentum of any of these particles by trying to move faster than them (i.e., perform a *boost*), since massless particles travel at the speed of light, and we can't outrun something moving at the speed of light. This means that in the early Universe, there was no difference between a left-handed electron and a left-handed neutrino. Even more curiously the right-handed electrons and neutrinos do not transform under the electroweak $SU(2)$ at all. This means that our Universe is not *invariant* under the interchange of left and right, aka parity, or in other words a mirror version of our Universe would have different physical effects. It was quite surprising when this odd behavior was revealed in the 1950s. Similarly the left-handed up and down quarks form an electroweak $SU(2)$ spinor, as do the left-handed top and bottom quarks. Because of this preference of left-handed fermions, the electroweak group is sometimes referred to as $SU(2)_L$. Needless to say, currently there is a very big difference between an electron and a neutrino; they have different electric charges and different masses. In our current understanding, it is the Higgs field that is responsible for this lack of symmetry; it is a consequence of the Higgs field having a nonzero average value now. To see how this come about, we need to think about the energy associated with different values of the Higgs field.

In order to compare the energy of different Higgs field values, we need to use $SU(2)_L$ *invariants* made out of the Higgs field, since the energy does not change under an $SU(2)_L$ transformation. We also need to restrict ourselves to real (rather than complex) *invariants*, since the energy is a real number, not a complex number. As we have seen, there is a real number associated with every complex number $z = x + iy$, that is, its modulus

$$|z| = \sqrt{x^2 + y^2}. \tag{12.148}$$

It is helpful to use the complex conjugate of z in order to calculate the modulus. The complex conjugate of

$$z = x + i\,y = r\,e^{i\theta} \tag{12.149}$$

is found by taking i to $-i$, so the complex conjugate of z is

$$z^* = x - i\,y = r\,e^{-i\theta} \tag{12.150}$$

and the square of the modulus is

$$|z|^2 = z^* z = (x - i\,y)(x + i\,y) = x^2 + y^2 = r\,e^{-i\theta} r\,e^{i\theta} = r^2. \tag{12.151}$$

If we write the Higgs field spinor as

$$\begin{pmatrix} H_1 \\ H_2 \end{pmatrix}, \tag{12.152}$$

then we can form another spinor out of its conjugate:

$$\begin{pmatrix} -H_2^* \\ H_1^* \end{pmatrix}. \tag{12.153}$$

So we can make a real $SU(2)_L$ *invariant* from

$$(-H_2^*, H_1^*) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = H_1^* H_1 + H_2^* H_2 = |H_1|^2 + |H_2|^2. \tag{12.154}$$

Now in the standard model of particle physics, the potential energy density for the Higgs field is given by

$$V(H) = \lambda \left(|H_1|^2 + |H_2|^2 - \frac{v^2}{2} \right)^2, \tag{12.155}$$

where $v = 246\,\text{GeV}$. The potential energy density $V(H)$ gives the energy per unit volume for a Higgs field that has the same value everywhere in spacetime; to find the total energy in a certain volume, we would multiply the potential energy density by the volume. If there are spatial variations in the Higgs field, there is an additional energy that needs to be added.

As we have seen this potential energy density looks have a maximum when the Higgs field vanishes and a minimum at a nonzero value (see Fig. 12.7). We can take the Higgs field at the minimum to be

$$H = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \tag{12.156}$$

whether v appears in the bottom or the top of the spinor is simply a convention that we fix by ensuring that the nonzero component has no electric charge.

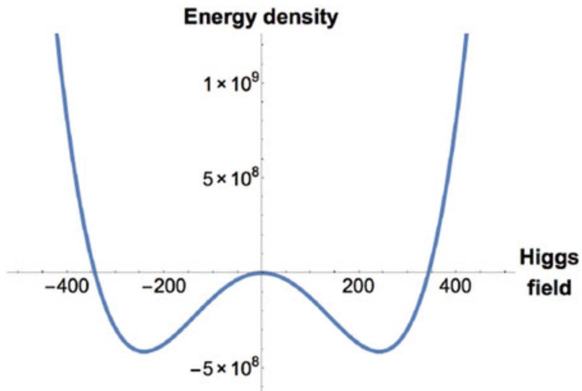


Fig. 12.7 The potential energy density of the Higgs field, $V(H)$, in the standard model of particle physics.

Let's look in more detail at what happens when the Higgs field is near the average value shown in Eq. (12.156). Let's write the Higgs field in terms of some small fluctuations, h, g_0, g_1, and g_2 around the average value:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} g_1 + i\, g_2 \\ v + h + i g_0 \end{pmatrix}. \tag{12.157}$$

Then the Higgs quadratic *invariant* is given by

$$|H_1|^2 + |H_2|^2 = \frac{1}{2}\left(g_1^2 + g_2^2 + (v+h)^2 + g_0^2\right). \tag{12.158}$$

Putting this into the potential energy density (12.155), we have

$$\begin{aligned} V(H) &= \frac{\lambda}{4}\left(g_1^2 + g_2^2 + (v+h)^2 - v^2 + g_0^2\right)^2 \\ &= \frac{\lambda}{4}\left(g_1^2 + g_2^2 + 2vh + h^2 + g_0^2\right)^2. \end{aligned} \tag{12.159}$$

Focusing on just the quadratic terms (the terms with only two powers of fields), we find

$$V_{\text{quad}} = \lambda v^2 h^2. \tag{12.160}$$

This tells us that one of the fluctuations, h, has a mass, while the remaining fields, g_0, g_1, and g_2, have no mass terms. The fact that these fields have no mass terms in the potential is not an accident, it is a consequence of the potential energy density being *invariant* under $SU(2)_L$ rotations, while the solution for the minimum of the potential, given in Eq. (12.156), is not *invariant*. The symmetry argument for the absence of mass terms in the potential energy density is very general and can be applied in many other cases. The general argument is known as Goldstone's theorem. We will return to these three special, massless fields a little later.

The fluctuations in the field h are what we identify with the Higgs boson that was discovered at the Large Hadron Collider. The Higgs boson mass is given by

$$m_h = \sqrt{\lambda}\, v = 125\,\mathrm{GeV}, \tag{12.161}$$

so we see that the self-coupling of the Higgs field, λ, is crucially involved in determining the Higgs boson mass.

When the temperature of the expanding Universe cools sufficiently, the Higgs can minimize its energy by taking on a nonzero value. While a Higgs field of $H = (0, 0)$ is *invariant* under $SU(2)_L$ transformations, a nonzero value is not. Having a nonzero value for the Higgs field breaks the $SU(2)_L$ symmetry. Now we can understand how particles can get masses. In the early Universe, top and bottom quarks had to be massless in order to preserve the symmetry between the left-handed top quark and the left-handed bottom quark. But we can include extra energy terms that involve both left-handed and right-handed quark fields and the Higgs field. We will write the electroweak top-bottom quark spinor as

$$\begin{pmatrix} t_L \\ b_L \end{pmatrix} \tag{12.162}$$

and call the right-handed top and bottom quark fields t_R and b_R. Then we can write an energy term that is $SU(2)_L$ *invariant*

$$\begin{aligned} y_t\, t_R^*(t_L, b_L) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} &= -y_t\, t_R^* t_R H_2 t_L \\ &\quad + y_t\, t_R^* b_L H_1, \end{aligned} \tag{12.163}$$

where y_t is the strength of the interaction between the Higgs field and top quarks. We need to add the complex conjugate of this term as well in order to have a real energy density. Even though this term respects the $SU(2)_L$ invariance of the theory, when the Higgs has a nonzero average value, we see a term that looks like

$$-\, y_t\, \frac{v}{\sqrt{2}} t_R^* t_L \tag{12.164}$$

which means that a left-handed top quark can turn into a right-handed top or in other words the top quark has a mass of

$$m_t = y_t\, \frac{v}{\sqrt{2}}. \tag{12.165}$$

We can also add a term

$$\begin{aligned} & -y_b\, b_R^*(t_L, b_L) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -H_2^* \\ H_1^* \end{pmatrix} \\ & = y_b\, b_R^* t_L H_1^* - y_b\, b_R^* b_L H_2^*, \end{aligned} \tag{12.166}$$

so when $H_2 = v/\sqrt{2}$, we see a mass for the bottom quark:

$$m_b = y_b\, \frac{v}{\sqrt{2}}. \tag{12.167}$$

The fact that the formerly indistinguishable top and bottom quarks have different masses clearly shows that the $SU(2)_L$ symmetry is broken.

▷▷▷ **More Accurate !!** Writing the interaction terms for spin one half particles is a little more involved than I am showing here, since you also need to keep track of the spin state of the particle, which involves another spinor index.

The value of the Higgs field also plays an important role in changing the massless gauge bosons of the electroweak interactions into the massive W^+, W^-, and Z bosons of the weak interactions and the massless photon of electromagnetism. Massless spin-one particles like the photon have two polarizations (aka degrees of freedom). That is why a single polarizer blocks half of an unpolarized light beam while a second polarizer rotated 90° will block all of the remaining lights. However, massive spin-one particles like the W's and Z must have three polarizations, which means that the Higgs must also produce the three extra polarizations required to give them masses. As we have seen, since the Higgs field is a complex valued, two-component spinor, it has four degrees of freedom, that is, each of the two-spinor components has a real and imaginary part. When the Higgs field is at the minimum of its potential energy, three of these degrees of freedom—g_0, g_1, and g_2—join with three of the electroweak gauge bosons to form the massive W^+, W^-, and Z. That leaves one degree of freedom left over, which is just h, the piece that corresponds to the Higgs boson. The masses of the W^+, W^-, and Z are also proportional to the value of the Higgs field, v, and to their coupling to the Higgs field, which is in fact the "gauge coupling." The photon doesn't couple directly to the Higgs field and that is why it remains massless. The result that the three seemingly massless fields, g_0, g_1, and g_2, end up as part of the massive spin-one fields is known as the "Higgs mechanism."

The standard model of particle physics makes detailed predictions about all the properties of the Higgs boson. For example, we can easily calculate the Higgs boson lifetime. The Higgs couples directly to other particles proportionate to their mass, so the biggest quantum amplitudes are for the Higgs boson spontaneously transforming into the heaviest possible particles in the standard model. However, in order to conserve energy, the final two particle masses must add up to less than the Higgs boson mass, so with $m_H = 125\,\text{GeV}$, the Higgs boson is too light to decay into a top quark and an anti-top quark or two Z bosons or two W bosons. Thus, the Higgs boson mostly decays to a bottom quark and an anti-bottom quark (Fig. 12.8).

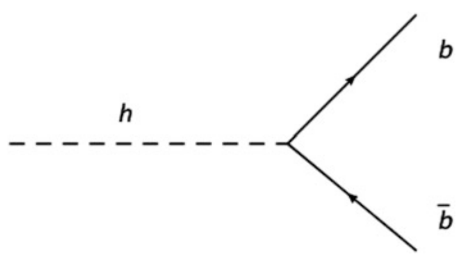


Fig. 12.8 The decay of the Higgs boson into a bottom and anti-bottom quark, time is moving to the *right* in this diagram.

The spin averaged, squared probability amplitude for a Higgs boson to turn into a bottom quark and an anti-bottom quark is

$$|\mathcal{M}|^2 = \frac{2N_c\, m_b^2}{v^2}\left(m_H^2 - 4\, m_b^2\right), \tag{12.168}$$

where I've written the coupling y_b in terms of v and m_b, and the factor $N_c = 3$ reflects that fact that quarks come in three possible "colors" of the strong interaction. In the reference frame where the Higgs is sitting still, each quark carries away half the Higgs rest energy $E_H = m_H c^2$, so

$$E_b = \frac{m_H c^2}{2} = \gamma m_b c^2, \tag{12.169}$$

From this we can check that the quarks have momentum

$$p_f = \frac{m_H}{2}\left(1 - \frac{4\, m_b^2}{m_H^2}\right)^{1/2}. \tag{12.170}$$

Using Eq. (12.65), and integrating over spherical angles (which merely gives a factor of 4π), we find

$$\mathit{\Gamma}_{\text{Higgs}} = \frac{p_f}{32\,\pi^2\, m_H^2}\int_0^{2\pi}\int_0^{\pi}|\mathcal{M}|^2 \sin\theta d\theta\, d\phi = \frac{N_c\, m_b^2\, m_H}{8\pi\, v^2}\left(1 - \frac{4\, m_b^2}{m_H^2}\right)^{3/2}. \tag{12.171}$$

Taking the bottom quark mass as $4.18\,\text{GeV}/c^2$, this gives $\mathit{\Gamma} \approx 4.3\,\text{MeV}/\hbar$, and a lifetime

$$t_{\text{Higgs}} = \frac{1}{\mathit{\Gamma}_{\text{Higgs}}} \approx 1.5 \times 10^{-22}\,\text{s}, \tag{12.172}$$

a very short lifetime indeed!

A curious feature of including quantum effects in the interactions of particles is that coupling constants aren't constant. This is not just a silly version of Murphy's law but a real, measurable effect. When we measure the strength of electromagnetic interactions at different length scales, we find that the coupling is stronger at shorter distances. For electromagnetism we usually describe the strength of the interaction in terms of the "fine-structure constant" α. When we measure α on atomic scales, we find

$$\alpha(r = 10^{-10}\,\text{m}) \approx \frac{1}{137}. \tag{12.173}$$

Many people spent way too much time trying to come up with a numerological reason for the number 137 appearing in the denominator here. This turned out to be a waste of time, because when we measure α at a much smaller distance, we find

$$\alpha(r = 2 \times 10^{-18}\,\text{m}) \approx \frac{1}{128}. \tag{12.174}$$

The point is that while in a classical world with $\hbar = 0$, we would see that α was a constant; in our world there are energy-dependent quantum corrections. For example, if an electron exchanges a photon with another charged particle, there are many things that can happen to the photon as it propagates between the two charged particles. Two possibilities are shown in Fig. 12.9.

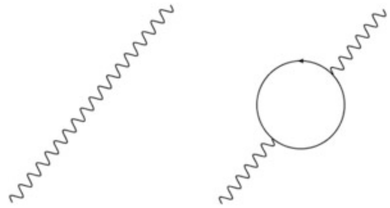


Fig. 12.9 The classical ($\hbar = 0$) contribution to photon propagation is shown on the *left*: nothing happens to the photon as it travels. On the *right* we see a quantum loop correction to photon propagation: the photon briefly turns into an electron-positron pair, which then turn back into a photon that continues on its way. Time is moving to the *right* in this diagram.

Quantum mechanics allows the photon to briefly turn into an electron and a positron and then back into a photon. This process is energy dependent, since it only makes a significant contribution when the photon energy is larger than the rest energy of the electron, $m_e c^2$. Higher photon energies correspond to higher photon momentum, and these are the photons that dominate in short distance processes. If we have a sufficiently energetic photon, we can have any charged particle-antiparticle pair in the loop. Typically these processes give α a logarithmic dependence on energy. The change in the strength of couplings with scale is called "renormalization group running."

The same kind of effects also contributes to the self-coupling of the Higgs field, which we called λ. Figure 12.10 shows two contributions to the scattering of two Higgs particles. The classical contribution, proportional to λ, just corresponds a term in the potential energy density (12.159) with four powers of Higgs fields. The quantum correction contributes a logarithmic dependence on energy. There are many other processes that can also contribute.

In the approximation where we only keep the classical term and the term involving the single top quark loop, we find that the strength of the coupling measured at a high-energy scale $\mu\, c^2$ is related to the strength of the coupling measured at an energy scale corresponding to the top quark mass $m_t\, c^2$ by

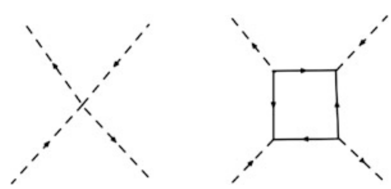


Fig. 12.10 The classical contribution and quantum top quark loop contribution to the four-Higgs interaction.

$$\lambda(\mu\, c^2) = \lambda(m_t\, c^2) - \frac{4N_c\, y_t^4}{16\pi^2} \ln\left(\frac{\mu}{m_t}\right), \tag{12.175}$$

where $\ln(x)$ is the natural logarithm. Eventually, if we go to high enough energies, the second term will overcome the first, and $\lambda(\mu\, c^2)$ will become negative. This is what we saw happening in Fig. 9; when we are studying very large values of the Higgs field, we need to use the self-coupling evaluated at a correspondingly high-energy scale, and so eventually the potential starts to decrease, and we have hit the rim of the volcano. Needless to say, in order to do this calculation properly, we need to take many different kinds of quantum corrections into account that is more or less straightforward. The harder part is that we need to know all of the possible particles that can appear in the loops at very high energies that requires some strong assumptions. People who do these calculations assume that there are no new particles beyond the ones we already know, that is, they are assuming that the standard model of particle physics is correct all the way up to enormously high energies that we haven't probed experimentally. If there are some heavier new particles to be discovered, then the calculation would have to be completely changed.

12.11 Fusion

Inside the Sun the process of fusion mainly proceeds by converting four protons into a helium nucleus (an alpha particle), two positrons, and two neutrinos and releasing excess energy as photons. The first step in this process involves two protons (which we will write as p^+) converting to deuterium (a nucleus with a proton and a neutron, written as D^+) as well as a positron (e^+) and a neutrino (ν). We can write this deuterium production reaction as

$$p^+ + p^+ \to D^+ + e^+ + \nu. \tag{12.176}$$

The positron will quickly meet an electron and they will annihilate into two photons, while the neutrino will quickly leave the Sun, since it only experiences weak interactions involving the heavy W and Z bosons, and stream away into space. The deuterium can combine with another proton to produce tritium (a nucleus with two protons and one neutron) and a photon. Finally two tritiums can convert to an alpha and two protons.

The first step in this process, deuterium production, is the limiting step, because it is far slower than the others; this is because it involves the exchange of a heavy W boson. We can write the *cross section* for this process in terms of the kinetic energy, K, associated with the relative velocity, v, of the protons,

$$K = \frac{1}{2}\mu v^2, \tag{12.177}$$

where μ is the "reduced mass." If we have two particles with masses m_1 and m_2, then the "reduced mass" is

$$\mu = \frac{m_1 m_2}{m_1 + m_2}, \tag{12.178}$$

so in our case $\mu = m_p/2$. The *cross section* for deuterium production (for energies at or below the typical energies of particles in the Sun) is

$$\sigma(p^+ + p^+ \to D^+ + e^+ + \nu) \approx \frac{4\times 10^{-22}\,\mathrm{keV\,b}}{K} \exp\left(-2\pi\alpha\sqrt{\frac{m_p c^2}{2K}}\right) \tag{12.179}$$

The numerical factor on the right-hand side of Eq. (12.179) is so small because it is suppressed by four powers of the ratio of the deuterium binding energy (about 2 MeV) to the W boson rest energy ($m_W c^2 \approx 80\,\mathrm{GeV}$). The reason it is suppressed by the W boson mass is that in this process a proton must convert to a neutron. At the quark level, this means that an up quark converts to a down quark. An up quark can convert to a down quark by emitting a virtual W boson. The W boson is virtual because the energy available is much smaller than its rest energy, so the process is highly suppressed. This suppression is what gave the weak interactions their name. The exponential factor in Eq. (12.179) reflects the fact that if the velocity is small (which means the kinetic energy is small), then the protons do not get close together (see Fig. 10.3), and there is a very small probability for deuterium to form unless the protons are within 10^{-15} meters of each other. This is because the proton and neutron are bound together by the strong interactions to form deuterium, but the strong interactions produce a short range force which does not extend past 10^{-15} meters.

For any system that has a temperature T, Boltzmann showed that states with a range of different energies can be present and that the probability of a particular energy state being present depends exponentially on the energy. For example, the probability of a proton having a kinetic energy K is

$$P(K) = A\,\exp\left(-\frac{K}{k_B\,T}\right), \tag{12.180}$$

where A is a constant that depends on how many protons are present and k_B is Boltzmann's constant. This means that particles having very high kinetic energies compared to $k_B T$ are very unlikely. Using this distribution Boltzmann also showed that the typical kinetic energy of any particle is given by

$$K_{\mathrm{typical}} \approx \frac{3}{2}\,k_B T, \tag{12.181}$$

where the factor of 3 reflects the fact that the particles could be moving in any of the three different directions in space.

Since the probability distribution (12.180) is falling exponentially with kinetic energy, while the *cross section* is rising exponentially with kinetic energy, there is a balancing point: a kinetic energy where the reaction is most likely to occur. It turns out that the deuterium production process is most likely to occur at the Gamow energy

$$K_G = (\pi\,\alpha\,k_B T)^{2/3}\left(\frac{1}{4}m_p c^2\right)^{1/3} \tag{12.182}$$

In order to calculate the rate at which the reaction actually proceeds, we need to average the *cross section* times the relative velocity over the Boltzmann distribution (12.180), because there is always a range of possible energies that are allowed. This average is called the thermal average, which we will write by putting angle brackets, $\langle\ \rangle$, around the quantity that was averaged.

One finds that

$$\langle \sigma\, v\rangle = 4\times 10^{-22}\,\text{keV b}\left(\tfrac{32}{3\,m_p}\right)^{1/2}\left(\tfrac{2K_G}{(k_BT)^4}\right)^{1/6} \exp\left(-\tfrac{3K_G}{k_BT}\right). \tag{12.183}$$

So at the center of the Sun, where the temperature is 1.5×10^7 Kelvin, the typical kinetic energy is 1.9 keV and the Gamow energy is 7.4 keV, and the deuterium production *cross section* is

$$\sigma(p^+ + p^+ \to D^+ + e^+ + \nu) \approx 5\times 10^{-31}\,\text{b}. \tag{12.184}$$

The rate of fusion depends on the number density of protons, $n_p = 4.45\times 10^{31}\,\text{m}^{-3}$, at the center of the Sun. If we look at an individual proton, the number of fusion reactions per second would be

$$\Gamma_{\text{fusion}} = n_p\,\langle\sigma v\rangle = 2.6\times 10^{-18}\,\text{per second} \tag{12.185}$$

which corresponds to a mean time for fusion to occur (per proton) of

$$t_{\text{fusion}} = \frac{1}{\Gamma} = 3.9\times 10^{17}\,\text{seconds} = 1.2\times 10^{10}\,\text{years} \tag{12.186}$$

99 % of the power generated in the Sun is produced inside the "solar core," a sphere with a radius of $0.24\,R_\odot$ that is about a quarter of the Sun's radius. The volume of the solar core is

$$V = \frac{4}{3}\pi\,(0.24\,R_\odot)^3 = 1.9\times 10^{25}\,\text{m}^3, \tag{12.187}$$

so total rate of deuterium production in the Sun is

$$R = \frac{1}{2}n_p^2\langle\sigma v\rangle\,V = 1.1\times 10^{39}\,\text{per second} \tag{12.188}$$

where we have multiplied by one half to avoid double counting the two identical protons in the initial state of the fusion reaction (12.176).

About 3.6×10^{38} protons are converted to helium per second in the core of the Sun, which means that 4.3×10^9 kg of mass are converted to energy per second, which provides a power output of 3.8×10^{26} Watts.

Further Reading

R.P. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton University Press, Princeton, 1985)

G.F. Giudice, *A Zeptospace Odyssey: A Journey into the Physics of the LHC* (Oxford University Press, Oxford, 2010)

L. Hoddeson et al. (ed.), *The Rise of the Standard Model: Particle Physics in the 1960s and 1970s* (Cambridge University Press, Cambridge, 1997)

L.M. Krauss, *Hiding in the Mirror: the Mysterious Allure of Extra Dimensions* (Viking, New York, 2005)

L.M. Krauss, *The Greatest Story Ever Told—So Far: Why Are We Here?* (Atria Books, New York, 2017)

L. Randall, *Warped Passages: Unravelling the Mystery of the Universe's Hidden Dimensions* (HarperPerennial, New York, 2006)

L. Randall, *Knocking on Heaven's Door: How Physics and Scientific Thinking Illuminate the Universe and the Modern World* (Ecco, New York, 2011)

More Technical Reading

E.G. Adelberger et al., Solar fusion cross sections, https://arxiv.org/pdf/astro-ph/9805121v2.pdf

N. Kaloper, J. Terning, How black holes form in high energy collisions, https://arxiv.org/pdf/0705.0408.pdf

Image Credits

Figure 7.1: John Terning / CC BY-NC-SA 2.0.

Figure 7.2: John Terning / CC BY-NC-SA 2.0.

Figure 7.3: John Terning / CC BY-NC-SA 2.0.

Figure 7.4: John Terning / CC BY-NC-SA 2.0.

Figure 7.5: John Terning / CC BY-NC-SA 2.0.

Figure 7.6: from *Las Maravillas Del Cielo* by Don Rogue Galvez Y Encinar, 1876 / Public Domain.

Painting of Thomas Young: portrait by Sir Thomas Lawrence / Public Domain.

Figure 7.7: user:Jordgette / Wikimedia Commons / CC-BY-SA-3.0 / GFDL.

Figure 7.8: John Terning / CC BY-NC-SA 2.0.

Figure 7.9: Frank Trixler, Ludwig-Maximilians-Universität München.

Photo of Ernest Rutherford: by George Grantham Bain in his collection at the Library of Congress / Public Domain.

Drawing of ATLAS detector: Argonne National Laboratory / Wikimedia Commons / CC-BY-SA-2.0.

Figure 7.10: mashup by User:Zykure / Wikimedia Commons / CC-BY-SA.

Figure 8.1: user: Raylet2 / Youtube http://bit.ly/2eSQOJA.

Photo of Max Planck from 1878: / Wikimedia Commons / Public Domain.

Painting of Michael Faraday: oil, by Thomas Phillips,1842 / Public Domain.

Figure 8.2: John Terning / CC BY-NC-SA 2.0.

Relatvisitic rockets: John Terning / CC BY-NC-SA 2.0.

Figure 8.3: John Terning / CC BY-NC-SA 2.0.

Figure 8.4: National Aeronautics and Space Administration / Public Domain.

Rutherford's drawing: Ernest Rutherford, in *Radioactive transformations*, 1906, http://bit.ly/2eHDNmP / Public Domain.

Figure 9.1: John Terning / CC BY-NC-SA 2.0.

Figure 9.2: John Terning / CC BY-NC-SA 2.0.

Figure 9.3: John Terning / CC BY-NC-SA 2.0.

Figure 9.4: Heather Rowley, http://www.rowleysoapworks.com.

Figure 9.5: John Terning / CC BY-NC-SA 2.0.

Figure 9.6: John Terning / CC BY-NC-SA 2.0.

Figure 9.7: John Terning / CC BY-NC-SA 2.0.

Figure 9.8: John Terning / CC BY-NC-SA 2.0.

Sphere visiting Flatland: John Terning / CC BY-NC-SA 2.0.

Parallel 2-branes: John Terning / CC BY-NC-SA 2.0.

Spectrum of electromagnetic waves: John Terning / CC BY-NC-SA 2.0.

Figure 10.1: John Terning / CC BY-NC-SA 2.0.

Figure 10.2: adapted from Michael H. Repacholi et. al., *Bioelectromagnetics* Volume 33, Issue 3, pages 187–206, April 2012, DOI: 10.1002/bem.20716.

Figure 10.3: John Terning / CC BY-NC-SA 2.0.

Figure 12.1: John Terning / CC BY-NC-SA 2.0.

Figure 12.2 adapted from user JabberWok / English language Wikipedia / CC-BY-SA-3.0.

Figure 12.3: adapted from user JabberWok / Wikimedia Commons / CC-BY-SA-3.0.

Figure 12.4: adapted from user Inductiveload / Wikimedia Commons / Public domain.

Figure 12.5: John Terning / CC BY-NC-SA 2.0.

Figure 12.6: adapted from user Jakob Scholbach / Wikimedia Commons / CC-BY-SA-3.0.

Figure 12.7: John Terning / CC BY-NC-SA 2.0.

Figure 12.8: John Terning / CC BY-NC-SA 2.0.

Figure 12.9: John Terning / CC BY-NC-SA 2.0.

Figure 12.10: John Terning / CC BY-NC-SA 2.0.

Part III: Misconceptions about the Multiverse

Yasunori Nomura

Introduction

The multiverse refers to a picture discussed in current theoretical physics, which says that what we once thought as the entire universe is only one of the many universes in which physical laws take different forms. There are lots of misconceptions in public about this picture, which we will discuss in this part. Common misconceptions include "The multiverse is not a scientific, or scientifically motivated, theory"; "Adopting the anthropic principle, implied by the multiverse picture, is equivalent to giving up scientific explanations for any phenomena we observe"; and "The multiverse theory is not testable even in principle." We will see that all these statements are false; in particular, we will see that the approach of multiverse cosmology is exactly that of conventional science.

Chapter 13

The Basic Picture

In this section, we overview the multiverse theory. We will see that it is a concrete proposal based on fundamental theories of physics. In particular, we will see that it is scientifically well motivated, both observationally and theoretically.

13.1 Definition

What does the end of the universe look like—what exists outside of it? How was the universe born—what was there before it was born?

To answer these questions, we first need to define what we mean by the "universe." For example, suppose there is some "end" of the universe, and there is "something" outside of it. But if we decide to call everything including that something as the universe, then *by definition* there would be no such thing as the outside of the universe. Since we will be talking about a *concrete* picture, we need to define things precisely.

We know from the era of Copernicus and Galilei that the land we live on is only the surface of one of the planets in the solar system. We also now know that our solar system is only one of many such objects in our Milky Way galaxy, which in turn is only one of many galaxies. If we keep

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_13

Our universe is almost homogeneous when viewed at large scales.

looking at larger scales in this way, we find that our universe is roughly homogeneous (after "coarse-graining" structures at smaller scales). Furthermore, this homogeneous region is well described by the so-called standard model of particle physics, more precisely the standard model extended to include what are called dark matter and dark energy. We call this homogeneous region, described by a single particle physics model, our universe.

One of the major discoveries in the twentieth century is that this region, which we call the universe, is expanding. (What was actually found is that the more distant a galaxy is, the faster it recedes from us, which implies that the universe is expanding as a whole.) When you hear that the universe is expanding, you might imagine that the size of the universe is finite and increasing in time. This is, however, not necessarily the case.

▷▷▷ **Misconception !!** The expanding universe means a finite universe expanding in ambient space.

The precise meaning of the universe being expanding is, as shown in Fig. 13.1, that the distance between arbitrary chosen two points keeps increasing as time passes. In particular, the "size" of the universe can be infinite *from the beginning*, except possibly at the exact time zero when the distance between any points can be said zero or undefined. (We usually consider what happens at the exact time zero to be a result of mathematical idealization anyway.)

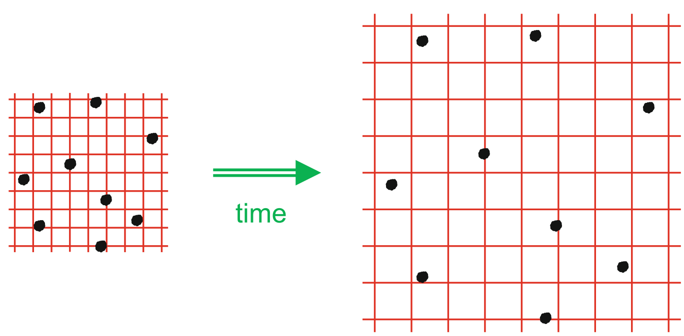


Fig. 13.1 In the expanding universe, the distance between arbitrary two points increases as time passes.

⊳⊳⊳ **More Accurate !!** The size of the expanding universe can be infinite all the way through its history.

The fact that the universe is expanding—in the sense described above—means that at early times it was much more dense and hence at higher temperature. Namely, the universe must have started from a hot "big bang" phase. In fact, the universe in the early hot phase was observed directly. Since the speed of light is finite, seeing a distant object (e.g., at a thousand light years away) means that we are seeing the light emitted from the object in the past (a thousand years ago). So, if we keep looking farther and farther, at some point we must see a hot, high temperature and high density, universe. This implies that the background night sky must be shining! On the other hand, the night sky is obviously dark. What is going on?

The answer is that because the universe is expanding, the light from the hot phase is subject to the Doppler effect, so that its frequency is in the microwave region when received on the Earth. In other words, when viewed by microwave, the night sky is shining. This radiation, coming uniformly from all the angles in the sky, is called the cosmic microwave background (CMB)—it is light emitted when the universe was $400,000$ years old (which is only $1/35,000$ of the current age of the universe, 13.8 billion years); see Fig. 13.2.

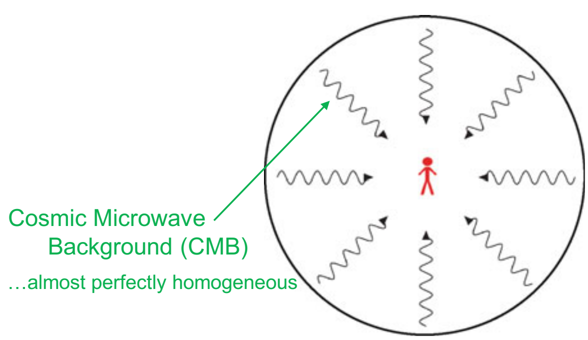


Fig. 13.2 We receive the cosmic microwave background (CMB) from all the directions in the sky, which is almost perfectly homogeneous.

Following the initial discovery in 1964, the CMB has been measured in many ground-based and satellite experiments. A striking fact is that these measurements tell us that the universe, when it was young (400,000 years old), was incredibly homogeneous. In fact, the fluctuations of energy density at that time were only at the level of one part in 100,000. All the structures we see today—stars, galaxies, galaxy clusters, and so on—have been formed because these tiny fluctuations were amplified by "gravitational instability." Namely, a region slightly more dense than the surroundings becomes more dense as gravity attracts more matter to that region; accordingly, regions that were less dense have become even less dense. We can indeed use the measured map of the CMB to simulate what happens to the universe afterward, and the result well agrees with the observed structure of galaxies and galaxy clusters in the current universe.

According to the measurement of the CMB, our universe was almost perfectly homogeneous when it was younger.

The universe before 400,000 years old cannot be seen directly by light, since the high density of the universe prevents any light from propagating. We can, however, extrapolate the history of the universe further back, using the equations of general relativity and the standard model of particle physics. (The current age of the universe—13.8 billion years old—is obtained in this manner.) Through this, we know what happened in the earlier universe, for example, how light chemical elements were synthesized when the age of the universe was only about 1–10 minutes—the predicted relative abundances of the elements agree well with the observation. The history before this "big-bang nucleosynthesis" era was not fully settled observationally, but we have a rough idea: at some early time, the universe was subject to exponential expansion called inflation (which provided the origin of small fluctuations needed to form structures), after which an asymmetry between the amounts of matter and antimatter was created through the process called baryogenesis (whose details are still debated). In any event, it seems clear that the early universe was very homogeneous, i.e., it looked very much the same everywhere.

Let us consider drawing this history of the universe in a figure which looks "scientific." In doing so, physicists often draw a "spacetime" figure in which a spatial (time) direction is taken to be in the horizontal (vertical) direction and in which the trajectory of a light ray is taken to be in a

45° direction. (This type of figures is called Penrose diagrams, and it makes causal relations between events manifest. For example, since no signal can propagate faster than the speed of light, the spacetime region a point in the figure can affect is that between the upper left and upper right directed 45° lines drawn from that point.) According to this rule, (the history of) the universe we have been discussing so far, naively, seems to be summarized as in Fig. 13.3.

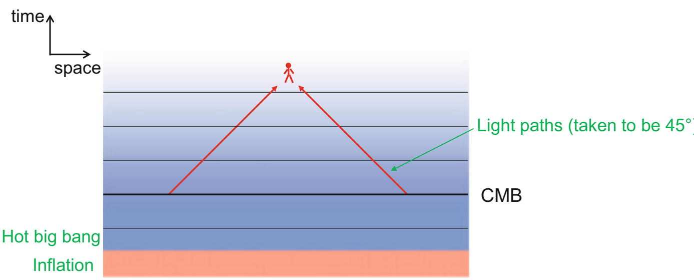


Fig. 13.3 The naive picture of the universe in the form of a Penrose diagram.

What the multiverse theory says is that this innocently-looking figure is wrong, at least in one important way. In particular, if we look at Fig. 13.3, we might conclude that there is no room to consider other universes spatially separated from our own. This is, however, not correct. As we will see, according to the multiverse theory, the structure of spacetime is more intricate than that in Fig. 13.3, so there can be other universes even in a region spatially separated from us (in some specific sense; see below).

▷▷▷ **Misconception !!** The multiverse is a vague idea just suggesting that there are many universes.

▷▷▷ **More Accurate !!** The multiverse, as discussed currently in the theoretical physics community, is a specific proposal for the structure of space and time, replacing the naive picture of the universe in Fig. 13.3.

13.2 Motivation: Observational

To go beyond the naive picture of Fig. 13.3, we need some hints. Below we will discuss some of them, which played important roles in the development of the multiverse picture. In doing so, we will address the following misconception:

▷▷▷ **Misconception !!** The multiverse is a random guess which is not scientifically well motivated.

In particular, we will see that the multiverse is well motivated both observationally and theoretically.

Our universe seems to be carefully designed so that complex structures, including intelligent life, can emerge.

One of the greatest mysteries of our universe is that it appears that it is designed too well for us humans. For example, let us imagine changing the value of the mass-squared parameter μ^2 of the Higgs field in the standard model, which can theoretically take any value in the range $-10^{32}\mu_0^2 \lesssim \mu^2 \lesssim 10^{32}\mu_0^2$, where μ_0^2 is the value in our universe. We then find that unless μ^2 is in an extremely narrow range, $0 \lesssim \mu^2 \lesssim$ a few $\times\, \mu_0^2$, there is no stable nucleus (except for hydrogen, which is the proton). Namely, our universe acquires enough complexity to have nuclear physics, and hence chemistry and life as we know, only if this parameter is carefully chosen to be in this tiny window. We also find that the masses of various elementary particles (which are determined by certain coupling constants in the standard model) must also be chosen carefully in order for the theory to possess complexity. These complexities are presumably a necessary condition for life to emerge. Who made such careful adjustments of the theory?

This issue became even more mysterious when experimental collaborations led by Perlmutter, Riess, and Schmidt discovered in 1998 that the expansion of the universe is accelerating, rather than decelerating. Under normal circumstances, the expansion of the universe is only decelerating because the gravitational force between any matter is attractive. However, the expansion can accelerate if space is filled by the "energy of the vacuum." The 1998 discovery, therefore, implies that our universe is filled with the vacuum energy (or at least something that effectively behaves as the vacuum energy).

What surprised people was that the size of the observed vacuum energy density was extremely small—120 orders of magnitude smaller than theoretically expected! And yet it was nonzero. Theoretically, the vacuum energy density ρ_Λ is expected to take any value in the range $-\rho_{\Lambda,*} \lesssim \rho_\Lambda \lesssim \rho_{\Lambda,*}$, where $\rho_{\Lambda,*} \sim 10^{90}\,\mathrm{g/cm^3}$ represents a theoretically expected size, and yet the actual value found by the observations is surprisingly close to zero, $\rho_\Lambda \sim 10^{-120}\rho_{\Lambda,*}$. Moreover, the observed value of ρ_Λ is very special—it is only about a factor of 2 different from the energy density of matter:

$$\rho_\Lambda \sim 2.2\,\rho_{\mathrm{matter}} \tag{13.1}$$

Note that the two could have taken values many orders of magnitude different, and yet they are this close. This becomes even more mysterious if we realize the fact that the time dependencies of the two quantities are completely different: as the universe expands, the energy density of matter is diluted inversely proportional to the volume, while the vacuum energy density stays constant. What we find is the fact that these two components are comparable in size *when humans make cosmological observations* (see Fig. 13.4). In other words, if we try to explain the smallness of the vacuum energy by some mechanism that operated in the early universe, then the mechanism must know when intelligent life—humans—will emerge and make cosmological observations, to adjust the vacuum energy density in such a way that it becomes comparable to the matter energy density at the time the observations are made. Can we imagine any such mechanism?

Our universe has comparable vacuum and matter energy densities now.

In 1987, 11 years before the discovery of accelerating expansion, Steven Weinberg published a paper about the vacuum energy in *Physical Review Letters*. The problem of smallness of the vacuum energy density was already known by then. Most physicists, however, were thinking that this was not a pressing problem. Their thinking was: given that the vacuum energy density was already smaller than its natural size by more than 100 orders of magnitude, its true value would be zero due to some unknown mechanism, and in any case the solution to this problem would not much affect the rest of physics.

Most physicists once thought that the vacuum energy in our universe is zero.

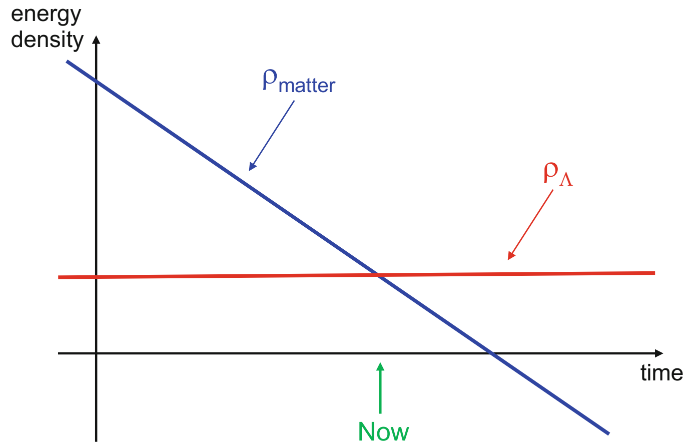


Fig. 13.4 We live in a very special era in which the energy densities of vacuum and matter are comparable. Here, the *horizontal* and *vertical* axes should be understood to represent the logarithms of time and energy densities.

Weinberg did not think that way. There were some suggestions for how the vacuum energy could be zero, but none of them was working well. (He wrote a review article, summarizing why these ideas do not work.) Instead of pursuing another such mechanism, he considered what would happen if the vacuum energy were much (e.g., by a few orders of magnitude) larger than the current matter energy density. He then found that in such a universe, there would be no structure such as galaxies (and hence any intelligent life). This implies the following. Suppose there are a large number of universes in which the vacuum energy takes different values. (In the discussion here, it is not important how these universes are realized.) Then, some of these universes would accidentally have a value of the vacuum energy small enough to lead to nontrivial structures. Since intelligent life would emerge only in such universes, when they make cosmological observations, they always find a surprisingly small value of the vacuum energy density; see Fig. 13.5.

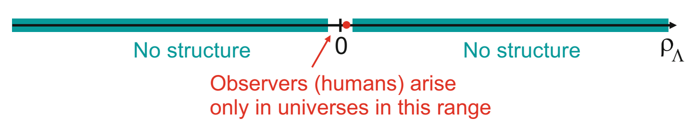


Fig. 13.5 Observers arise only in universes in which the vacuum energy density is sufficiently small.

An important point is that in this picture, which postulates many universes or the multiverse, the vacuum energy is expected to be *not much smaller than needed*. In other words, the size of the vacuum energy density is expected to be roughly comparable to the current matter energy density. This is in sharp contrast with what most physicists were imagining back then. And indeed, the observations by Perlmutter, Riess, and Schmidt found that the vacuum energy density is within only about a factor of 2 from the current matter energy density!

▷▷▷ **More Accurate !!** The multiverse scenario was considered in order to address a specific scientific problem that could not be solved by other theories. It not only explained the smallness of the vacuum energy but also predicted that it is comparable to the current matter energy density. This prediction was confirmed observationally a decade later.

Similarly, other "miraculous" features of the standard model that seem to be carefully designed for the existence of life can also be understood if we take the view that we live only in one of the many universes which satisfies the conditions for life to exist.

13.3 Motivation: Theoretical

It is certainly a major assumption that there are many possible universes in the world. Is there any independent argument supporting this hypothesis beyond the fact that it can explain the seemingly well-designed nature of our universe? In fact, there is.

Finding a complete quantum mechanical theory of gravity has been a challenging avenue. A problem is that when one straightforwardly quantizes Einstein's theory of general relativity, the resulting theory suffers from uncontrollable divergences at the fundamental level, suggesting that we need to do something more dramatic to have a consistent theory of quantum gravity. String theory is the leading contender for such a theory. In string theory, the extended nature of fundamental constituents tames these divergences. This is virtually the only quantum gravitational theory we currently have in our hands (although they are people who are exploring alternatives, such as loop quantum gravity).

String theory predicts the existence of extra spatial dimensions.

String theory predicts that the dimension of spacetime is higher than $four$, the number we see around (three spatial dimensions and one time dimension). In a certain way of counting, the number of extra dimensions is 6, and they are all spatial. What does it mean that there are extra spatial dimensions? Imagine that our world exists on the surface of a thin tube. This space is clearly two-dimensional at the fundamental level. However, if we are interested only in physics at large distance scales, e.g., because we are large, then this space appears as one-dimensional as illustrated in Fig. 13.6. So, for such "large" observers, the space appears

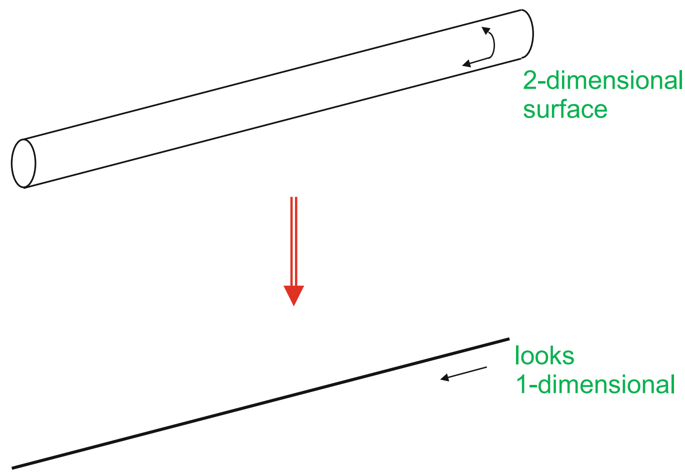


Fig. 13.6 The surface of a thin tube, which is two-dimensional, appears as one-dimensional at long distances.

as one-dimensional, with one extra dimension—in this case a circle—attached at each point in space. Similarly, in string theory, we must consider that six-dimensional space is attached at each point of our four-dimensional spacetime.

An important point is that the properties of the theory describing physics in four-dimensional spacetime at long distances depend on the shape and size of the extra dimensions because it results, in a sense, from averaging out the physics associated with these tiny compact dimensions. In particular, the content and properties of elementary particles and the value of the vacuum energy change if the shape and size of the extra dimensions are varied. (In fact, the number of "compactified" dimensions may also be other than six, so the number of spacetime dimensions at long distances may also differ from four.)

How many different configurations for the extra dimensions does the theory allow? It is quite common in nature that even if the fundamental equation is simple, a system governed by the equation shows remarkable complexity and varieties. A good example is organic macromolecules that all arise as quasi-stable solutions to a simple Schrödinger equation. Likewise, even though the equation governing the dynamics of the extra dimensions is simple, these dimensions can have an enormous number of quasi-stable configurations—an estimate says that there are 10^{500} or more such configurations. We can illustrate this situation schematically as in Fig. 13.7, where the horizontal directions represent possible configurations of the extra dimensions and the vertical axis corresponds to the potential energy associated with each configuration. Each quasi-stable configuration corresponds to a minimum of the potential valleys. This picture is often called the string landscape.

There are an enormous number of quasi-stable configurations for the extra dimensions.

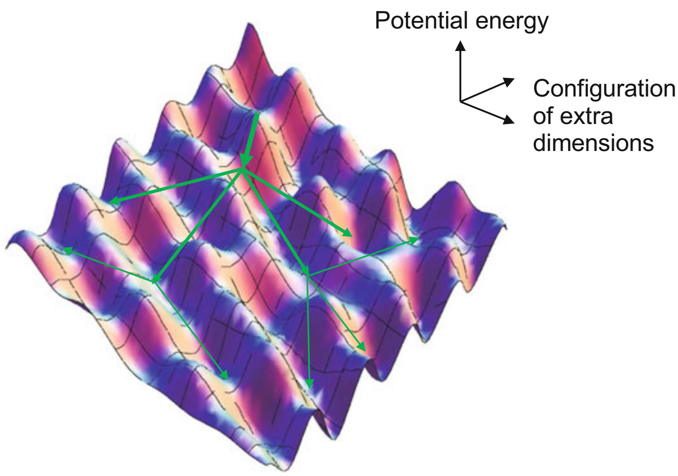


Fig. 13.7 A sketch of the potential energy as a function of the configuration (shape and size) of the extra dimensions. Each minimum corresponds to a different low-energy universe.

We conclude that each minimum of the above potential corresponds to a possible four-dimensional universe, having distinct elementary particles and the vacuum energy. However, this by itself is not sufficient to solve the problem of the vacuum energy, since it merely says that string theory has many universe as possible *solutions*. In order to lead to the multiverse and hence for Weinberg's solution to the problem to operate, these universes must be *physically realized*. What does the theory say about this?

Suppose the system was originally at some minimum of the landscape potential which has a positive potential energy. In this case, Einstein's equation tells us that spacetime expands exponentially. The system, however, does not simply stay there forever. Quantum mechanics allows a transition from that minimum to a lower minimum through a quantum tunneling process. Under a normal circumstance, such a quantum tunneling process occurs in the following manner. Initially, the system is everywhere in the state of a higher minimum. At some point, however, small bubbles form in which the system is in the state of a lower minimum (like bubbles in boiling water). These bubbles then expand almost at the speed of light, and they collide with each other, eventually turning the entire system into the new state in the lower minimum of the potential.

However, in the case of cosmic tunneling under consideration, ambient space in which bubbles form is expanding exponentially, in fact at a rate faster than the speed of bubble expansion. The bubbles, therefore, cannot fill the entire space—there is always ambient space exponentially expanding, in which new bubbles keep forming. Bubbles formed in this way can be of various different kinds: interiors of these bubbles correspond to different minima in the potential landscape, as indicated by the arrows in Fig. 13.7. In fact, it is generally expected that all the different universes in string theory are physically realized as bubble universes in this manner. This process of eternally creating bubble universes in exponentially expanding ambient space is called the eternal inflating multiverse.

Different universes predicted in string theory are physically realized as bubble universes through eternal inflation.

This setup, therefore, is exactly the one needed by Weinberg to solve the problem of small vacuum energy. According to this picture, our universe is only one of (infinitely) the many bubble universes formed through eternal inflation; see Fig. 13.8 for illustration.

Interestingly, the fact that string theory can lead to a huge number of low energy theories and that inflationary expansion at a high potential energy minimum lasts eternally was known from the 1980s, but people viewed them as nuisances. In particular, many people thought that the existence of extra dimensions in string theory is an unfortunate feature of the theory (as they thought it would have been better if it predicted four dimensions), and the eternal nature of inflation is a problem, since in our universe such

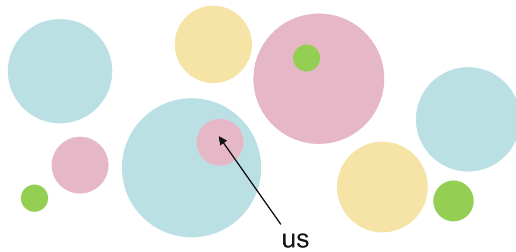


Fig. 13.8 Eternally inflating space produces an infinite number of bubble universes, one of which is our own universe.

expansion certainly ended before the hot big bang era. It is quite suggestive that properties of the fundamental theories which once people thought as bad features are exactly those needed to understand aspects of our universe as we observe it today.

▷▷▷ **More Accurate !!** The multiverse picture is suggested by the fundamental theories—string theory and eternal inflation—that were developed independently of the multiverse. It is, therefore, theoretically well motivated.

13.4 New Spacetime Picture

According to the multiverse theory, we live in one of the bubble universes which is nucleated in inflating space and expanding afterward. This may sound like our universe is not homogeneous even approximately—in particular, there is a special point corresponding to the center of the bubble, and if we go far in the bubble, then we would hit its wall, i.e., the edge. Doesn't this contradict the fact that our universe appears observationally very homogeneous as we saw in Sect. 13.1?

In fact, two statements that (i) we live in a bubble universe which was born small in ambient space and expanding afterward and (ii) our universe is (almost perfectly) homogeneous do not contradict with each other. How is this possible?

It is often the case that when a revolutionary change of the picture occurs, it is accompanied by the corresponding change of a concept. For example, when ancient people realized the possibility that our land might be spherical, one of the strongest "scientific" objections was that if it is indeed spherical, then people in the other side would "fall to the sky." Of course, we now know what is wrong with this argument—the concept of "down" is not universal to everyone; it is defined (only) with respect to the Earth through gravitational attraction. Similarly, in the multiverse picture, we need to embrace some revision of a concept, which makes the two statements above consistent.

There is no absolute concept of equal time at spatially separated points.

What concept do we need to revise among those we take for granted in our daily life? The point here is to realize that there is no *absolute* definition of equal time at spatially separated points. For example, there is no invariant meaning to the question "what is going on at some specific point in the Andromeda galaxy when you are reading this sentence," since we cannot uniquely determine what time in the Andromeda galaxy corresponds to "now" here. You might think that we can define equal time in two different places if we prepare two clocks synchronized at some location and then move them to the two places, e.g., one here and one in the Andromeda galaxy. But because of relativistic effects, such a definition depends on how we carry the clocks, e.g., the path and speed of the transportation.

This fact plays a crucial role in understanding the structure of a bubble universe. Let us imagine that we are seeing a bubble universe from outside. Then, the universe is born small and then becomes larger in ambient space. If we write this in a Penrose diagram, in which a trajectory of light is drawn as a 45° line, it becomes as in Fig. 13.9. Here, the

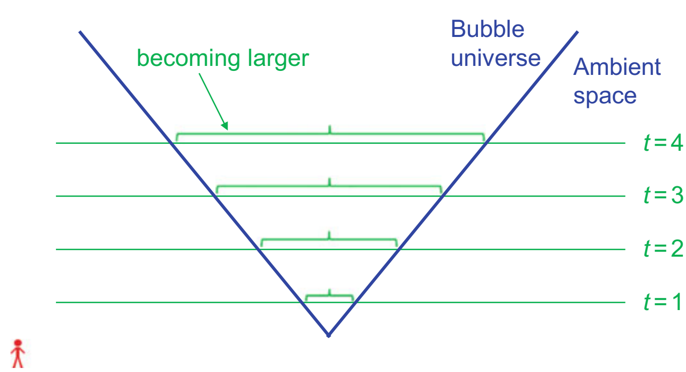


Fig. 13.9 Nucleation of a bubble universe as viewed from an exterior observer.

region inside (outside) the inverse triangle represents the interior (exterior) of the bubble universe. (The reason why the boarder of the two regions is given by 45° lines is that the bubble wall expands almost at the speed of light.) The equal time slices as viewed from the exterior observer are drawn as horizontal lines, denoted as $t = 1, 2, 3$, and 4. One can see that the size of the universe (the portions of the horizontal lines inside the inverse triangle) becomes larger as time passes.

On the other hand, if we see the same bubble nucleation process from the viewpoint of an observer inside the bubble, then it appears quite differently. In this case, equal time slices are given as in Fig. 13.10, denoted by $t' = 0, 1, 2, \cdots$. (Technically, equal time slices as viewed from an interior

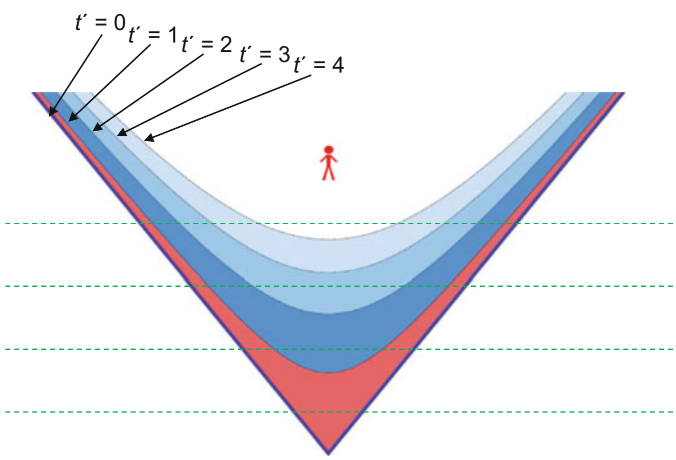


Fig. 13.10 A bubble universe as viewed from an interior observer.

observer are given by contours of quantum fields responsible for the nucleation.) We find that in this view, the universe is infinitely large already when it was born, i.e., the length of any constant t' line is infinite. Moreover, the universe, as viewed from the interior, is completely homogeneous, i.e., any point on a given constant t' line looks the same as any other point on the same line. (For $t' = 0$, there appears to be a special point—the vertex of the inverse triangle. This is an artifact of the drawing. Indeed, for any

$t' > 0$—how small it is—there is no special point on the equal time surface.)

When viewed from outside, a bubble universe is born small, and its size becomes larger as time passes. When the same universe is viewed from the interior, however, it is homogeneous and infinitely large already at the time it is born.

This is how the two statements about a bubble universe—that it is born small and becomes larger and that it is homogeneous—can be compatible. The former is a statement when the universe is viewed from outside, and the latter is that when viewed from the interior.

This picture replaces the naive picture given by Fig. 13.3. When our universe is drawn in the form of a Penrose diagram, it must be drawn as in Fig. 13.9 or Fig. 13.10, instead of Fig. 13.3. The new figure makes it clear that the spacetime region an exterior observer describes as "outside the universe" is the region "before the universe began" for an interior observer. And this is the region where other universes, born in the ambient eternally inflating space, as well as the ambient space itself reside. The full picture of the multiverse drawn in a Penrose diagram, therefore, is given by Fig. 13.11. One finds that it exhibits a "fractal" structure. Note, however, that many universes drawn near the top are in fact large universes because the entire spacetime is expanding. The fact that these universes appear small is an artifact of the rule of Penrose diagrams in which the trajectories of light are drawn as 45° lines.

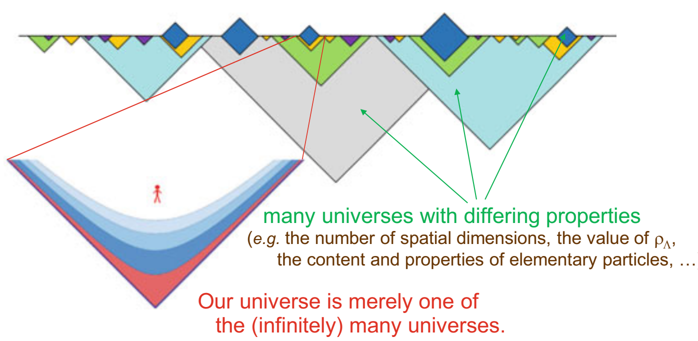


Fig. 13.11 A Penrose diagrammatic depiction of the multiverse.

Chapter 14

Demystifying the Multiverse

In this section we discuss some misconceptions about the multiverse, often existing even in the scientific community.

14.1 Scientific and Conservative

People sometimes say that the multiverse theory sounds "mystical." This is probably because it talks about a very big picture such as outside of our own universe.

▷▷▷ **Misconception !!** The multiverse picture, talking about things like outside of our universe, is mystical.

The true situation, however, is the opposite.

Suppose there was only one universe. Then it would be very difficult to explain miraculous features of our universe, such as the structure of elementary particles and the value of the vacuum energy, without resorting to some sort of creator. In the multiverse picture, however, there are an enormous number (10^{500} or more) of different universes, so some of them possess these miraculous features that lead to intelligent life, without a help of any creator. This, of course, does not prove that there is no such creator, but

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_14

given that a goal of science is to try to understand our physical nature as much as possible without relying on such an almighty person, the approach of the multiverse is exactly that of science.

▷▷▷ **More Accurate !!** The approach of the multiverse theory is exactly that of science.

In fact, the logic that has led to the multiverse picture is a very traditional one in science. A progress in science, especially in physics, often occurs in the following steps. First, to explain known facts, a new equation—or theory—is written down. Then, by studying that equation, we find new phenomena that were not known before. Finally, by accumulating evidence for these new phenomena, we build up our confidence about the equation, and through this process the new theory is becoming a part of our established scientific knowledge.

Two well-known historical examples are the following (see Fig. 14.1). The first is a story about relativity and gravity. When Einstein presented special relativity in 1905, this theory was not compatible with the known theory of gravity. (In relativity no signal can propagate faster than the speed of light, but gravity in Newton's theory propagates instantaneously.) This problem was solved when Einstein replaced Newton's theory with his theory of general relativity in 1916. This new theory—mathematically represented by the so-called Einstein equation—predicted new

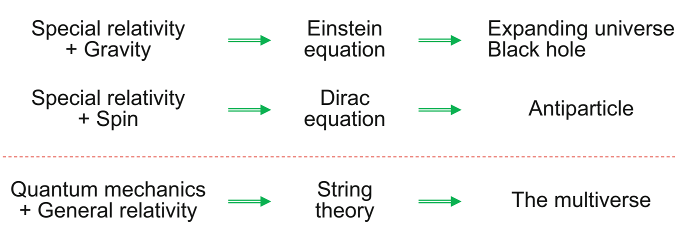


Fig. 14.1 The development of the multiverse theory in analogy with other scientific theories.

phenomena such as the expanding universe and black holes, which were not known at the time but confirmed later observationally.

The second example is when Paul Dirac tried to find an equation that allowed for describing a particle with spin in a way consistent with special relativity. The equation he found—the Dirac equation—predicted the existence of an antiparticle: for any particle there is a partner that has the same mass but the opposite charge. In Dirac's case, he was interested in writing the equation for the electron, so this antiparticle was what is now known as the positron, which was discovered 4 years after Dirac wrote down his equation.

An interesting thing is that in both these cases, those who originally wrote down the equations could not accept these predictions initially. It is well known that Einstein could not accept the prediction that the universe must be expanding (or contracting); the prejudice that the universe must be static was so strong back then. Dirac also tried to identify the positively charged particle predicted by his equation as the proton—the only known positively charged particle at the time—despite the fact that the equation was telling that the mass of the particle was the same as the electron. Even with the geniuses of Einstein and Dirac, it was not easy to overcome prejudices deeply insinuated in people's mind. And in both these cases, the equations once written down were describing the nature correctly beyond their originators' imagination.

The situation in the multiverse is not so different from these two examples (see Fig. 14.1). String theory has been considered as the virtually unique candidate for reconciling quantum mechanics and general relativity. Studying the structure of this theory has led to extra dimensions and hence many different universes at long distances. As in the case of the above two examples, this prediction had not been taken too seriously in the 1980s among people who originally developed string theory. But the discovery of the universe's accelerating expansion in 1998 made people pay attention to this implication of the equation.

▷▷▷ **More Accurate !!** The progress of the multiverse theory is very much analogous to those of other scientific theories.

In fact, one can even say that the multiverse is a "conservative" picture in the sense that it is more along the lines of the past progresses in science. When string theory turned out to be a consistent theory of quantum gravity in the 1980s, many physicists thought that all the aspects of our universe—including the masses of all the elementary particles and the value of the vacuum energy—could be derived merely by solving its equation, and this picture dominated the particle physics community for two decades afterward. Such a picture, however, is much more "radical"—we have never reached such an ideal situation in our history of science.

Throughout the history of science, we learned many times over that we were much tinier substances than we had previously thought and that we do not in any sense occupy a central position in the physical world. In ancient times, we thought we lived on the unique, disklike world, but we now know that we live on one of the eight planets around the Sun, which is only one of a few hundred billion stars in our galaxy, which is in turn one of many galaxies in our observable universe. Given this, it does not seem unreasonable—or even seem natural—that what we considered the whole universe is actually only a small portion of some larger structure.

Given the history of science, the multiverse is a conservative picture.

14.2 Anthropic "Principle"

The reasoning used by Weinberg to address the smallness of the vacuum energy is often called (the weak form of) the anthropic principle. This term, however, is misleading. When you hear the phrase anthropic principle, you might imagine that we are introducing some new principle that has somehow to do with humans. However, once we admit that there are many different universes, it is nothing more than a statement of consistent logic. What we want to explain is the fact that "when we made observation, we found a small value of the vacuum energy." To see if this is consistent with a theory, there is no point in discussing the value of the vacuum energy without taking observers into account. In other words, there is absolutely no problem if the value of the vacuum energy is large in universes in which there is no observer to measure it.

With sufficiently many universes, the anthropic reasoning is nothing more than a statement of consistent logic.

A statement that is typically made to the anthropic reasoning is the following:

▷▷▷ **Misconception !!** Once we admit the anthropic principle, we can explain everything with it, so it is equivalent to giving up a scientific explanation for any phenomenon; in particular, the anthropic principle means that there is no point in searching for any mechanism for explaining natural phenomena.

This statement is wrong in many respects. First, it is wrong that the anthropic reasoning can explain *everything* by itself. For example, the standard model of particle physics contains a quantity called the θ parameter, which controls the size of the electric dipole moments of elementary particles. This parameter is known to be smaller than its theoretically expected size by more than ten orders of magnitude. The origin of this smallness, however, cannot be explained by the anthropic reasoning alone, since we can show that even if this parameter was larger than the current experimental upper bound by many orders of magnitude, there would virtually be no effect for the structure of our universe. This implies that there must be some mechanism (other than the simple anthropic reasoning) that is making this parameter small.

Another reason for why the statement quoted above is incorrect is that the anthropic principle by itself does *not* mean that there is no conventional mechanism to explain the structure of a theory. As seen in Sect. 13.2, the standard model has another parameter that is much smaller than the theoretically expected value: the Higgs mass-squared parameter μ^2. The origin of this smallness appears to be anthropic because there is no complex chemistry if this parameter takes a value slightly different from the observed one. On the other hand, there are many mechanisms/theories considered in particle physics which explain the smallness of μ^2, most notably weak scale supersymmetry. Suppose that the anthropic reasoning for the smallness of μ^2 is correct. Does it mean that a mechanism explaining the smallness of μ^2 is absent in our universe?

It doesn't. In the multiverse, intelligent life emerges only in universes in which μ^2 is sufficiently small to accommodate complex structures. Some of these universes have small μ^2 "accidentally" (as in the case of the vacuum energy), while some have it through a mechanism such as weak scale supersymmetry. In this case, which type of universe we find ourselves in depends on how many each type of universes exists in the multiverse, or more precisely the relative probability of finding two types of universes given the condition that intelligent life exists. If the latter type dominates, we would find ourselves living in a universe in which the smallness of μ^2 is *realized* through a mechanism, even if its ultimate *reason* is anthropic. Until we become capable of calculating the relevant probability, we cannot conclude the existence or absence of a mechanism without using observations. (In the case of the vacuum energy, we could not find *any* mechanism making it small, so we believe that the vacuum energy is small "accidentally.")

▷▷▷ **More Accurate !!** The fact that some feature has an anthropic origin does not mean that there is no mechanism leading to that feature. The existence or absence of a mechanism is determined by the statistics in the multiverse (which we still cannot calculate from the first principle).

Chapter 15

Relation to Observation

As discussed in Sect. 14.1, the final stage of developing a scientific theory is to compare it with observations. What is the situation of the multiverse about this?

First, it should be emphasized that a small but nonzero vacuum energy density was a prediction of the multiverse, which was confirmed observationally in 1998. As discussed before, most physicists thought that the vacuum energy density was smaller than its natural size by more than 120 orders of magnitude because it was zero. One can say that one of the predictions of the multiverse theory was already tested observationally.

A prediction of the multiverse, i.e., a small but nonzero vacuum energy density, was already tested observationally.

Are there any other tests of the multiverse? A common misconception is:

▷▷▷ **Misconception !!** Since we cannot physically go to other universes, the multiverse theory can never be tested, and hence is not scientific.

This is incorrect in several respects. First, as the diagram in Fig. 13.9 or Fig. 13.10 shows, the fact that we cannot go to the region outside our universe (outside the inverse triangle) does not mean that we cannot obtain a signal from that region. Note that we cannot even go to a distant galaxy or an era in which dinosaurs lived, but studying these subjects certainly belongs to the realm of science.

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7_15

Another point is that to test a scientific theory, we need not confirm *all* of its predictions. Indeed, even for "well-established" theories like quantum mechanics and the standard model of particle physics, not all the predictions have been experimentally tested—in fact, it would be impossible to do such a thing. In the case of the multiverse, the relevant question is what this theory predicts for things we can observe within our own universe. (A small but nonzero vacuum energy density is one of such predictions.)

The multiverse predicts that our universe has negative curvature.

In this respect, the multiverse theory based on string theory and eternal inflation as discussed here makes an important prediction: our universe must have negative spatial curvature. What is the curvature of the universe? Imagine that we take three points in the universe and construct a triangle by connecting these points by the "shortest paths." Here, the shortest paths can be considered to be the paths light rays would follow. In this case, it is not guaranteed in general that the three inner angles add up to 180° as we learned in middle schools. Even on a two-dimensional surface, the sum of the angles is larger (smaller) than 180° if the surface is, e.g., a sphere (has a saddle shape). We call space in which the sum of the inner angles of a triangle is larger (smaller) than 180° positively (negatively) curved space.

As seen in Fig. 13.10, the multiverse theory tells us that our universe is only one of many bubble universes, and for observers living inside it, the equal time slices are given as in constant t' lines in the figure. One can then show mathematically that such equal time slicing makes the space look negatively curved. Namely, the sum of the inner angles of a large triangle drawn in our universe must be always smaller than 180°!

▷▷▷ **More Accurate !!** The multiverse makes predictions for observables that can be measured in our universe, and hence the theory is testable.

The multiverse theory, however, cannot predict how much the space in our universe is curved. This leads to the following possible future scenarios (see Fig. 15.1). The measurement of the curvature of the universe is expected to improve by about two orders of magnitude in the next cou-

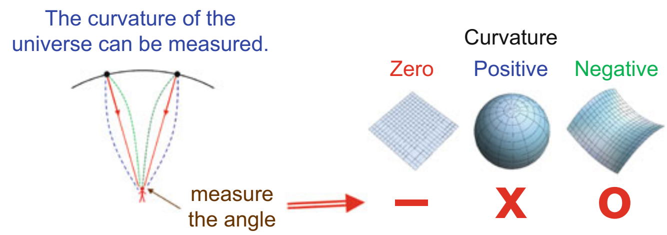


Fig. 15.1 The multiverse theory predicts that the spatial curvature of the universe is negative. If the future measurement finds the curvature to be (**a**) consistent with zero, (**b**) positive, and (**c**) negative, then the multiverse theory will be (**a**) unaffected, (**b**) excluded, and (**c**) supported, respectively.

ple of decades. The curvature of space can be found by measuring the angle between the two light rays emitted from an object whose size and distance from us are known (or calculated); see the left part of Fig. 15.1. Currently, a deviation of the sum of the inner angles of the largest triangle one can possibly draw in the universe is less than about a degree. The sensitivity to this deviation is expected to improve to the level of 0.01°.

Let us imagine that future measurements keep finding that the spatial curvature of the universe is consistent with zero. In this case, the result would be inconclusive for the multiverse. As mentioned above, the multiverse does not tell us how much the universe is curved—it depends on the length of slow-roll inflation that occurred in our bubble universe, which cannot (so far) be predicted from the first principle. This implies that the sum of the inner angles of a triangle may be, e.g., 179.999999999, in which case it is impossible to discriminate it from 180° observationally.

On the other hand, if a future measurement finds negative curvature, it would provide strong evidence for the multiverse. This is not only because the multiverse predicts negative curvature but also because it would be difficult to obtain a sufficiently small and yet measurable amount of curvature naturally, i.e., without fine-tuning, in a theory other than the multiverse. (In the multiverse, one can obtain a small but measurable amount of curvature quite naturally; for more details, see the paper quoted alongside.) If this happens, the multiverse theory would have another

For detailed discussions about the implications of the curvature measurement for the multiverse, see A.H. Guth and Y. Nomura, "What can the observation of nonzero curvature tell us?" Phys. Rev. D86 (2012) 023534 [arXiv:1203.6876 [hep-th]].

observational support in a way similar to the case of the small vacuum energy.

An interesting thing is that if a future measurement finds a positive nonzero curvature, then the multiverse as discussed here (based on string theory and eternal inflation) would be excluded. People often take falsifiability as a criterion for a good scientific theory. I myself think that we must be careful in applying this criterion idolatrously, since the judgment of falsifiability is often difficult, especially for a modern physical theory. But even if we adopt this criterion, the multiverse theory *is* falsifiable.

▷▷▷ **More Accurate !!** The multiverse theory is falsifiable.

If we are very lucky, we might even see more direct signals such as the remnant of a collision of our universe with another one. A Penrose diagram of this process is given in Fig. 15.2. As is clear from the figure, for an interior observer like ourselves, the effect appears as a signal from the spacetime region before the beginning of our universe; specifically, it appears as a slight deviation from homogeneity which existed already at the time the universe was born. The strength of the signal, however, is diluted by slow-roll inflation that has occurred within our bubble universe, so in order for the signal to be detectable, the length of slow-roll inflation must not be too long.

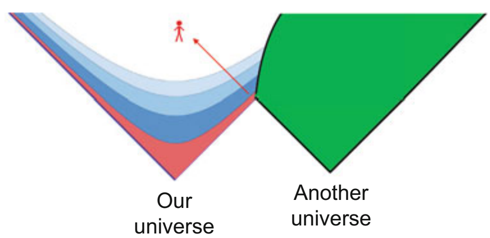


Fig. 15.2 A Penrose diagram representing a collision of our universe with another one.

As we have seen so far, the basic strategy to increase confidence about the multiverse is to accumulate observations that can naturally be explained by the multiverse picture but are difficult to obtain in any other way. Since we cannot physically go to other universes, this is clearly indirect, but this situation is not much different in some other cases, such as the big bang and inflationary theories. One might also be unsatisfied because the correspondence between the theory and observation is not "one to one": for example, even if we find a nonzero negative curvature in future observation, this measurement by itself would not uniquely select the multiverse as it is possible to come up with other theories that lead to such curvature even though the resulting theory would be contrived/fine-tuned. This situation, however, is also not so different from many other theories in the modern era. For example, in particle physics there is a theory called grand unified theory in which the three forces of the standard model are unified into a single force. It is often said that a smoking gun signature of this theory is that the proton decays (though with an incredibly long lifetime). Even in this case, however, it would not be impossible to come up with a model which leads to proton decay, although it may not be as elegant as grand unified theory.

The process in which we are observationally convinced with the multiverse theory could be frustratingly slow and likely would not be as dramatic as, e.g., a discovery of a new particle. However, here we are asking a particularly big question, and this is the cost we pay for it. At least, the framework is well motivated both observationally and theoretically to the extent we could probe so far. So it seems worth pursuing it further and see where it leads us to.

Epilogue

We have discussed misconceptions about the multiverse. A danger of these misconceptions is that they are sometimes used to argue against pursuing research on the subject. It is true that studying a big question involves risks, so it is legitimate if anyone decides personally not to participate in it. However, we never know how far a scientific theory can be developed until we seriously try to develop it. It seems fair to say that not everyone needs to study the multiverse, but someone should.

References and Further Reading

Here I list some selected textbooks and papers relevant to the subject discussed here. The list is in no sense exhaustive, but it gives an introduction to the vast number of articles written on the subject.

For the general subjects of early universe cosmology, string theory, and the anthropic principle, see, e.g.,

- J.D. Barrow, F.J. Tipler, *The Anthropic Cosmological Principle* (Oxford University Press, Oxford, 1988)
- K. Becker, M. Becker, J.H. Schwarz, *String Theory and M-Theory: A Modern Introduction* (Cambridge University Press, Cambridge, 2007)
- E.W. Kolb, M. S. Turner, *The Early Universe* (Westview Press, 1990)

and reference therein. For some key papers on the topics discussed here and on the multiverse in general, see, e.g.,

- V. Agrawal, S.M. Barr, J.F. Donoghue, D. Seckel, The anthropic principle and the mass scale of the standard model. Phys. Rev. **D57**, 5480 (1998) [hep-ph/9707380]
- R. Bousso, J. Polchinski, Quantization of four form fluxes and dynamical neutralization of the cosmological constant. JHEP **06**, 006 (2000) [hep-th/0004134].
- S. Coleman, F. De Luccia, Gravitational effects on and of vacuum decay. Phys. Rev. **D21**, 3305 (1980)
- A.H. Guth, E.J. Weinberg, Could the universe have recovered from a slow first order phase transition? Nucl. Phys. **B212**, 321 (1983)
- A.H. Guth, Y. Nomura, What can the observation of nonzero curvature tell us? Phys. Rev. **D86**, 023534 (2012) [arXiv:1203.6876 [hep-th]]
- S. Kachru, R. Kallosh, A. Linde, S.P. Trivedi, De Sitter vacua in string theory. Phys. Rev. **D68**, 046005 (2003) [hep-th/0301240].

- M. Kleban, Cosmic bubble collisions. Class. Quant. Grav. **28**, 204008 (2011) [arXiv:1107.2593 [astro-ph.CO]]
- A.D. Linde, Eternal chaotic inflation. Mod. Phys. Lett. **A1**, 81 (1986)
- Y. Nomura, Physical theories, eternal inflation, and the quantum universe. JHEP **11**, 063 (2011) [arXiv:1104.2324 [hep-th]]
- M. Tegmark, The multiverse hierarchy, in *Universe or Multiverse?*, ed. by B. Carr (Cambridge University Press, Cambridge, 2007), p. 99 [arXiv:0905.1283 [physics.pop-ph]]
- L. Susskind, The anthropic landscape of string theory, in *Universe or Multiverse?*, ed. by B. Carr (Cambridge University Press, Cambridge, 2007), p. 247 [hep-th/0302219]
- A. Vilenkin, The birth of inflationary universes. Phys. Rev. **D27**, 2848 (1983)
- S. Weinberg, Anthropic bound on the cosmological constant. Phys. Rev. Lett. **59**, 2607 (1987)

Index

A
Action, 69
Alpha particles, 91
Alpha rays, 105, 106
Andromeda galaxy, 190
Anthropic principle, 176, 196–198, 204
Antielectrons, 102
Atoms
  alpha particles, 91
  automatic electron detectors, 91
  carbon atoms, 90
  coherent light source, 87–88
  hydrogen atom, 83, 84
  Large Hadron Collider, 92, 93
  laser light, 88
  Lithium atom, 83
  microscope, 86–87
  photoelectric effect, 89
  position distribution, 84–86
  radio waves, 89
  visible light, 90
  wavelength, 88–89
  wave types, 89
  X-rays, 90

B
Bell-curve probability distribution, 28–29
Bell's inequality, 58
Bell's theorem, 48–49
Beta rays, 105, 106
Bipolar quantum
  characteristics, 15
  example, 15
  HUP
    attributes, position, and velocity, 21
    Born rule, 20
    classical observation, 22
    hidden variables, 22
    limiting principle, 20–21
    measurement issues, 20
  quantum probability waves
    conceptual and computational, 18–19
    observed position, 16–17
    wavefunction, 16–17
    wave of probability, 16
    wild-eyed new quantum theory, 18
  wavelike and particle-like attributes, 13–14
Black holes
  cosmic rays, 96
  dimensional analysis, 150
  Einstein's theory, 96, 98, 100
  escape velocity, 151
  evaporation process, 103
  extra dimensions, 153–155
  field lines, 99–101
  gravitational forces, 99
  Hawking's calculation, 102–103
  Hawking's evaporation, 152
  impact parameter,152
  Large Hadron Collider, 95, 96, 102
  lightweight, 102
  LIGO experiment, 97–98
  Newton's constant, 99
  Planck length, 98, 101–102, 150
  primordial black holes, 97
  Schwarzschild metric, 150
  Schwarzschild radius, 97, 98
  space-like direction, 151
Boltzmann's constant, 134, 169
Boson, 105–109, 123, 144, 164–166, 168, 169
Brane, 117
Bubble universes, 188–192

Y. Nomura et al., *Quantum Physics, Mini Black Holes, and the Multiverse*,
Multiversal Journeys, DOI 10.1007/978-3-319-41709-7

C
Cell phone radiation
  vs. brain gliomas, 122–123
  electromagnetic waves, 120, 121
  microwaves, 122
  sound waves, 121
Classical double-slit experiment, 29–30
Classical single-slit experiment, 27
Cold fusion, 80, 123–125
Compactified dimensions, 186
Complex numbers, 156
Configuration space, 37
Conjugate variables, 69–70
Copenhagen interpretation
  geocentric *vs.* heliocentric models, 63
  illusion/perception, 64
  many-worlds interpretation, 64–65
  new age cottage industry, 64
  objective collapse, 64
  wave-function collapse, 63–64
Cosmic microwave background (CMB), 179–181
Cosmic rays, 96
Cosmic tunneling, 188
Curved spacetime, 100, 148

D
Delayed choice experiments, 58
Dirac equation, 195
Double-slit experiment
  EPRB, 8, 9, 11
  soccer
    bell curves, 26
    classical experiment, 29–30
    example of, 26
    HUP, 26
    quantum experiment, 30–33
    single-slit experiment, 27–29
    tangible wave, 26
    wave interference, 26
    wave of probability, 25
Double-slit interference pattern, 70
Dynamical entanglement, 58

E
Einstein Podolsky Rosen Bell (EPRB) experiment
  American dimes, 51
  bulbs flashing, 53
  detector settings, 52–53
  experiment, 8, 9, 11
  identical coin detectors, 50
  quantum version, 54–55
  "tail/head" detectors, 51–52
Einstein Podolsky Rosen (EPR) experiment, 47
Einstein's special theory of relativity
  Anna's coordinates, 138
  arbitrary angle, 135
  boosted coordinate system, 138
  coordinate system, 135
  distances and times, 139–140
  events, 136–137
  invariant, 136, 140
  kinetic energy, 140
  origin, 136
  relativistic length contraction, 139
  rotating the map, 135
  spatial positions, 135
  Taylor series expansion, 140
  time interval to Bob, 139
Einstein's theory, 96, 98, 100
Electromagnetic waves, 120, 121
Electron scattering, 102
Entanglement, 5, 47–48, 58
Escape velocity, 97
Eternal inflation, 188–189, 200
Event horizon, 97
Expanding universe, 178
Extra dimensions, 98–103, 116, 117, 153–155, 186–188, 196

F
Fermions, 107
Force-carrying particles, 106
Fourier transform function, 70
Fusion
  deuterium production reaction, 168
  Gamow energy, 169–170
  probability distribution, 169–170
  reduced mass, 168
  relative velocity, 168
  solar core, 170
  thermal average, 170

G
Gamma rays, 105, 106
Gauge bosons, 106, 165
Generators of the rotation group, 157
Ghirardi, Rimini, and Weber (GRW), 64
Gluon, 106

The God particle, *see* Higgs boson
Goldstone's theorem, 163
Gravitational forces, 99
Gravitational potential energy, 108
Gravity
  coordinates of spacetime, 148
  Einstein's equivalence principle, 149
  escape velocity, 147, 151
  free-falling elevator, 149
  geodesic, 148
  gravitational potential energy, 146
  gravitational redshift, 149
  invariant interval, 148
  masses, 145–146
  spherical coordinates, 147–148
Group theory
  eigenstates, 160
  group of rotations, 159
  identity matrix, 158
  infinitesimal rotation, 157–159
  matrix multiplication, 157
  origin, 157
  quantum spins, 160
  rotation matrix, 158
  three-dimensional space, 157

H
Hadrons, 106
Hawking radiation, 103, 155
Heisenberg Uncertainty Principle (HUP)
  attributes, position, and velocity, 21
  Born rule, 20
  classical observation, 22
  double-slit experiment, 26
  hidden variables, 22
  limiting principle, 20–21
  measurement issues, 20
Helium atoms, 106
Heterotic theory, 116
Hidden variables, 5, 6, 22, 28, 45, 47–49, 54, 63
Higgs boson
  deflections, 107
  energy density, 109, 111
  force carriers, 106
  gluon, 106
  gravitational potential energy, 108
  hadron and lepton, 106
  quantum spin, 107
  quarks, 106
  rim, 110
  W and Z bosons, 109
Higgs field, 107–111, 156, 157, 161–165, 167, 168, 182
  anti-bottom quark, 165
  classical contribution, 167
  complex conjugate, 161–162
  electroweak group, 161
  energy-dependent quantum corrections, 166
  fine-structure constant, 166
  gauge coupling, 165
  Goldstone's theorem, 163
  Higgs boson mass, 164
  Higgs mechanism, 165
  invariants, 164
  massless spin-one particles, 165
  potential energy density, 162–163
  quantum correction, 167–168
  quark fields, 164
  renormalization group running, 167

I
Integrated luminosity, 155
Interpretations
  behavior and automatic thinking, 66
  causal narratives, 61
  causative agent, 67
  Copenhagen interpretation
    geocentric *vs.* heliocentric models, 63
    illusion/perception, 64
    many-worlds interpretation, 64–65
    new age cottage industry, 64
    objective collapse, 64
    wave-function collapse, 63–64
  depictions, 66
  metaphysics, 62
  treatments, 65–66

K
Kinetic-plus-potential energy, 70–71

L
Large Hadron Collider (LHC), 82–83, 95, 98, 101, 102, 106, 155
Leptons, 106
Lightweight black hole, 102
The LIGO, 97–98
Luminosity, 143, 155

M
Many-particle Schrödinger equation, 71
Mega electron-volts (MeV), 134
The Metric, 148, 149
Microwaves, 122
Milky Way galaxy, 187
Mini black hole, 95–103, 155
"M theory", 116, 117
Multiverse theory
  conservative, 196
  nonzero negative curvature, 203
  Penrose diagram, 202
  positive nonzero curvature, 202
  scientific theory, 193–195
  spatial curvature, 201

N
Negative probabilities, 58
Newton's constant, 99
Nonlocality, *see* Spooky action at a distance

O
Objective collapse, 64
Observer, 20, 25, 26, 37, 43–45, 64, 184, 190–192, 196, 200

P
Parallel universe
  brane, 117
  space and time coordinates, 112–115
  supergravity, 116
  symmetry transformations, 113
  10-dimensional theory, 116
Partial differential equation (PDE), 70
Particle colliders
  Higgs boson
    deflections, 107
    energy density, 109, 111
    force carriers, 106
    gluon, 106
    gravitational potential energy, 108
    hadron and lepton, 106
    quantum spin, 107
    quarks, 106
    rim, 110
    W and Z bosons, 109
  parallel universe
    brane, 117
    space and time coordinates, 112–115
    supergravity, 116
    symmetry transformations, 113
    10-dimensional theory, 116
Particles
  atoms
    alpha particles, 91
    automatic electron detectors, 91
    carbon atoms, 90
    coherent light source, 87–88
    hydrogen atom, 83, 84
    Large Hadron Collider, 92, 93
    laser light, 88
    Lithium atom, 83
    microscope, 86–87
    photoelectric effect, 89
    position distribution, 84–86
    radio waves, 89
    visible light, 90
    wavelength, 88–89
    wave types, 89
    X-rays, 90
  definition, 81–83
PDE, *see* Partial differential equation (PDE)
Photoelectric effect, 89
Planck length, 98, 101–102, 150, 154
Planck mass, 154
Planck's constant, 69
Planck time, 154
Position–momentum uncertainty principle, 69
Positive probabilities, 58
Positrons, 102, 123, 167, 168, 195
Primordial black holes, 97
Probability distribution function, 40
Probability wave, 16–19, 25, 26, 49

Q
Quantum double-slit experiment, 30–33
Quantum probability waves, 16–19, 25, 26
Quantum single-slit experiment, 28–29
Quantum spin, 107
Quantum wavefunction, 36
Quarks, 81, 82, 106
Quasi-stable configurations, 187

S
Scattering
  billiard balls, 141
  differential cross section, 143, 145
  energy and momentum conservation, 142

fine-structure constant, 144
fudge factor, 141
help-factor, 141
Higgs lifetime, 145
luminosity, 143
Planck's constant, 141
squared amplitude, 145
total cross section, 143–145
unbound electrons, 144
wave-length, 141–142
(Time-dependent) Schrödinger equation, 16, 71
Schrödinger's Cat, 8, 42–46
Schwarzschild metric, 150, 154
Schwarzschild radius, 97, 98
Scientific theory, 193–195
Single-slit experiment, 27–29
Singular quantum
accurate experimental predictions, 3
attributes, 5
debate, dissatisfaction, and disagreement, 4
EPR experiment, 5–6
experimental facts, 8–9
experts, 4–5
Feynman's credentials, 6
hidden variables, 6
intermediate-level audience, 10–11
misconceptions, 4
non-expert reader, 6–7
obvious fact, 4
ontological ambiguity, 3
preposterous fiction, 4
quantum consumer, 7–8
quantum-themed books, 7–8
string theory, 7
Spin attributes, 71
Spinor, 157–162, 164, 165
Spooky action at a distance
Bertlmann's socks, 48–49
classical probability, 49
EPR, 47
EPRB
American dimes, 51
bulbs flashing, 53
detector settings, 52–53
identical coin detectors, 50
quantum version, 54–55
"tail/head" detectors, 51–52
Grand Sudoku Challenge, 56–59
hidden variables, 49
measurement, 47–48
relativity theory, 48
Standard deviation, 69
Standard quantum theory, 38, 40
String theory, 7, 116, 185, 186, 188, 189, 195, 202, 204, 205
Supergravity, 116
Système international (SI) units, 69

T
Tangible waves, 31
10-dimensional theory, 116

U
Ultraviolet rays, 121
Uncertainty principle, 8, 20–23, 69, 70, 85, 102
Units
advantages, 131–132
Big Bang explosion, 131
dimensional analysis, 133
Einstein's equation, 132
four-dimensional spacetime, 134
Greek letters, 133, 134
Maxwell's innovations, 133
physics constants, 134
quantities, 131–132
Universe
Andromeda galaxy, 190
big-bang nucleosynthesis, 180
"big bang" phase, 179
bubble nucleation process, 191
bubble universes, 188–192
CMB, 179, 180
dark matter and dark energy, 178
definition, 17
Doppler effect, 179
Higgs field, 182
matter energy, 183, 184
Milky Way galaxy, 177
multiverse theory, 189
Penrose diagram, 181, 192
quantum gravitational theory, 185
quantum tunneling process, 188
quasi-stable configurations, 187
spacetime, 180–181, 186
string theory, 186, 187
vacuum energy, 182–185
Uranium nucleus, 124, 125

V

Vacuum energy density, 182–185

W

W and *Z* bosons, 109
Wavefunction
  many-particle, 36–37
  measurement
    absolute phase, 42
    collection of particles, 40
    complex numbers, 41
    constructive interference, 41
    data interpretation, 38
    experiments, 38–39
    relative phase, 41–42
    single-particle, 39
    standard quantum theory, 38
    trajectories, 42
  Schrödinger's cat experiment
    biomolecules, 46
    configuration space, 42–43
    macroscopic objects, 44, 46
    nano-soccer balls, 46
    nonlocality, 43
    particle interconnectedness, 43–44
    random collapse, 44
    wave of superposition, 44–45
Wavefunction collapse, 17, 63
Wave interference pattern, 31
Wave-particle duality, *see* Bipolar quantum
Web-based animations, 11

X

X-rays, 90

Z

Z bosons, 106, 109, 165, 168